1983

The transformation strain dependence on applied stress in equiatomic nickel-titanium alloys of powder metallurgical origin

Peter Ramon
Iowa State University

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Ramon, Peter

THE TRANSFORMATION STRAIN DEPENDENCE ON APPLIED STRESS IN EQUIATOMIC NICKEL-TITANIUM ALLOYS OF POWDER METALLURGICAL ORIGIN

Iowa State University PH.D. 1983

The transformation strain dependence on applied stress in equiatomic nickel-titanium alloys of powder metallurgical origin

by

Peter Ramon

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Nuclear Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1983

TABLE OF CONTENTS

ABSTRACT

The strain-temperature behavior of Ni-Ti samples upon martensitic phase transformation was investigated as a function of applied uniaxial tensile and compressive stresses. Axial and transverse strains were measured. The samples were fabricated from stock bars of powder metallurgical origin. They were machined into tubes or solid bars with square cross sections.

The tubular samples were cycled in temperature by circulating preheated or precooled silicone oil through them. At the same time, they were held in grips, which permitted a deadweight axial stress to be applied. Rosette-type strain gages were used to measure axial, circumferential, and 45°-inclined strains continuously as a function of temperature. The strain-temperature hysteresis loops were analyzed for the dependence of their shape (including magnitude and direction of transformation strain) on the type (tensile or compressive) and magnitude of the applied stress. Furthermore, the difference in behavior for Ni-Ti in the as-fabricated and annealed conditions was studied. In the absence of an applied stress, the strains upon transformation were nearly isotropic, and a net expansion took place upon transformation on cooling. By contrast, as the uniaxial tensile stress was increased, the axial and circumferential strains became increasingly different, i.e., the axial strain corresponded to an expansion upon transformation on cooling and the circumferential strain to a contraction. The magnitudes of these strains increased progressively with increasing stress, with the axial

strain magnitudes approximately twice those of the circumferential strains. As expected, the axial and circumferential directions were found to correspond closely to principal directions of strain. Similar results were obtained for uniaxial compressive stresses, except that the axial strain corresponded to a contraction and the circumferential strain to an expansion upon transformation on cooling.

The observation that the transformation strain magnitudes increase with increasing applied stress magnitudes is explained on the basis of a preferred activation model (PAM). The PAM incorporates the concept of stress-induced preferential activation of martensite variants, in contrast to the random activation of them in the absence of an applied stress. The application of the model to shape memory behavior contributes to a better understanding of the concept of training (the inducing of a transformation pattern as a result of prior cycling). It also explains the effect of loading sequence on the shape of the hysteresis loops. The phenomenological theory of Wechsler-Lieberman and Read (WLR) is applied to the case of martensitic phase transformation in Ni-Ti, and semi-axes of the total distortion ellipsoid are calculated. In conjunction with the PAM, the calculated total distortions predict the saturation limit of phase transformation strains with increasing applied uniaxial tensile stress.

For the solid square cross-sectional bar, single strain gages were bonded to all four sides. The strain gages permitted the axial and transverse strains to be determined upon temperature cycling for various

applied tensile stresses. The fractional change in volume was calculated as a function of temperature and stress from the measured strains. The net volume change upon transformation on cooling was found to be positive for all tensile stresses up to about 7 ksi.

1. INTRODUCTION

1.1 General Background

A wide class of materials, including metals, alloys, ceramics, minerals and inorganic compounds, exhibits first order, martensitic, phase transformations (MPT) from a parent phase, austenite, to a product phase, martensite [1]. The characteristics of an MPT on the basis of which it is distinguishable from other solid-state transformations are the following crystallographic relationships between the austenite and the martensite:

(a) The two sets of crystal axes have a unique orientation relationship that is governed by a coordinated lattice distortion (often designated the Bain distortion [2].) and certain rotations.

(b) A unique planar interface exists between the austenite and the martensite, called the habit plane, along which the macroscopic distortion strain is zero [3]. This condition requires an additional lattice invariant distortion, such as twinning.

If, in addition the transformation is thermoelastic [4], then it exhibits the property called shape memory effect (SME) [5].

The simplest SME (one-way memory) can be described as the ability of the material to regain its original shape at the high temperature phase (HTP), following a moderate deformation imposed on the material while in the low temperature phase (LTP) [6]. Thus, a piece of martensite memory

alloy which is slightly bent while in the LTP will "straighten out" if heated into the HTP regime. The reversion of the deformed martensite to the recovered austenite involves high recovery stress, thus suggesting the possibility of direct conversion of heat into mechanical work. Based on the availability of this work, one of the envisaged future uses of SME alloys is a low temperature solid-state engine [7].

Other possible technological applications may involve more or less diverse innovations, like hydraulic fittings [8], collapsing antennas [5], orthodontic dental arch wires [9], and actuating elements [10], to name just a few. The alloys exhibiting SME include AuCd [11], CuAlNi [12], CuZn [13], CuZnAl [14], AISI 304 stainless steel [15], NiTi [16], and InTl [17].

1.2 The Near Equiatomic Composition Nickel-Titanium Alloy

The intermetallic compound NiTi exists over a wide range of compositions, from 48 to 55 at.% Ni at high temperatures [18-21] as shown in the high temperature equilibrium phase diagram of Fig. 1 [22], but it has a restricted homogeneity range at low temperatures [23]. Further, it was found by Scholl et al., that the intermetallic compounds TiFe and TiCo are in solid solution with TiNi, provided the Ti concentration is constant at 50 at.% [24].

When examined by optical microscopy, the microstructure reveals the existence of finely dispersed secondary particles, regardless of the

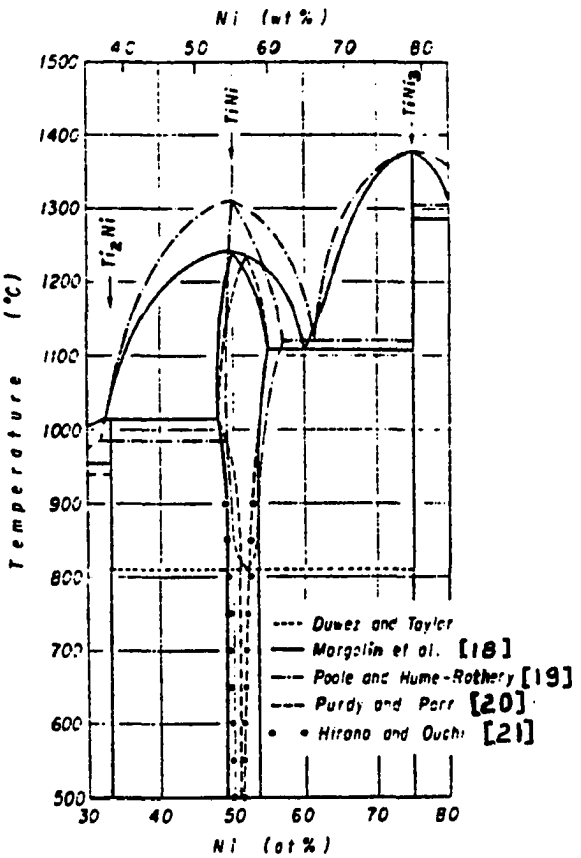


Figure 1. High temperature Ti-Ni equilibrium phase diagram in the vicinity of the NiTi phase [22].

source of the sample under investigation, whether laboratory samples, arc melted billets [25] or commercially prepared wires [26]. In many cases, the precipitates were found to be Ti_4Ni_2O, and were present even for relatively low oxygen levels. Because of the restricted homogeneity mentioned above, some Ti-rich Ti_2Ni phase, which has the face-centered cubic Fd3m crystal structure, may be stabilized by oxygen [27] to form Ti_4Ni_2O, with the same structure, but slightly larger lattice parameter than for Ti_2Ni [28]. In the nickel-rich region of the equilibrium phase diagram, the stable intermetallic compound is Ni_3Ti, with a hexagonal DO_{24} structure [29]. On photomicrographs, this nickel-rich compound may appear as Widmanstätten plates [27].

The crystal structure and lattice parameters of equiatomic Ni-Ti in the HTP and LTP have been studied using electron and x-ray diffraction techniques. It was found that the austenite is a single phase B2 type lattice (ordered BCC or CsCl structure) with a lattice parameter a_o=3.015Å [30-33], while the martensite is of distorted B19 lattice type, with a monoclinic unit cell, $P112_1/m$, the lattice parameters of which are: a=2.885Å, b=4.120Å, and c=4.622Å, with the angle between a and c axes being β=96.8° [33-34].

1.3 Phenomenological Theory, the Wechsler-Lieberman-Read Model (WLR)

One of the versions of the crystallographic theory of martensitic phase transformations is given by the WLR model [3]. The theory is general in that it does not depend upon experimental results specific to

a particular martensitic transformation system. The ideas underlying the theory are as follows: (a) There is a particular distortion (the lattice or Bain distortion) that brings about the essential change in crystal structure. For NiTi, this is the distortion that takes the B2 cubic lattice into the distorted B19 monoclinic lattice. Of all the distortions capable of doing this, the lattice distortion is the one that requires the smallest strains [35]. (b) In addition, there must also be (except in very special cases) a lattice invariant distortion, since the lattice distortion alone usually does not permit a plane of zero distortion to exist. The plane of zero distortion is generally considered to be necessary, since it must serve as the interface plane between the as-yet untransformed austenite and the already transformed martensite. Examples of the lattice invariant distortion are shears due to twinning, slip, or faulting. (c) The total distortion, combining (a) and (b), is then an invariant plane strain, that does permit the existence of a plane of zero distortion over macroscopic distances (i.e., distances covering many twin, slip or fault bands). This plane is identified as the habit plane on which the martensite forms.

In many cases, including NiTi [36], the martensite product is observed to be internally twinned. Let the total transformation distortion be represented by $\bar{\bar{E}}$. Then, a vector $\vec{r}$ is transformed to $\vec{r}'$, where

$$\vec{r}' = \bar{\bar{E}}\,\vec{r}$$

and where, for the twin-related Bain distortions $\bar{\bar{T}}_1$, and $\bar{\bar{T}}_2$, we have

$$\bar{\bar{E}} = \bar{\bar{\Phi}}_1[(1-x)\bar{\bar{T}}_1 + x\bar{\bar{\Phi}}\bar{\bar{T}}_2] \equiv \bar{\bar{\Phi}}_1\bar{\bar{F}} \qquad (2)$$

In equation (2), (1-x) and x are the fractional amounts of twins 1 and 2, respectively. Also, $\overline{\overline{\Phi}}$ is the relative rotation for twins 1 and 2, and $\overline{\overline{\Phi}}_1$ is the rotation for twin 1 alone. The theory determines x and ultimately $\overline{\overline{E}}$, such that a plane of zero macroscopic distortion exists. Once $\overline{\overline{E}}$ is determined, the crystallographic features of the transformation are evaluated. Thus, the habit plane, orientation relationships, and macroscopic distortion (plane and direction of the invariant plane strain) are obtained as a function of the lattice parameters of the two phases. The theory, as applied to the NiTi transformation, is given as an appendix. Of special interest for the present study, is the evaluation of the principal strains corresponding to the invariant plane strain for the macroscopic distortion.

1.4 The Self-Accommodating Nature of Martensite

The phenomenological theories are devoted to the consideration of crystallographic factors that attend the establishment of a plane of zero macroscopic distortion at the interface between the austenite and a single martensite plate. The present study will show that the transformation strains in a particular arbitrary direction are a strong function of the applied stress in that direction. The interpretative of this hinges on the idea that many crystallographically equivalent orientations or variants are possible. Wasilewski has pointed out [37] the importance of the self-accommodation of the various variants that may form in a given local region. Thus, the strains due to several individual plates should tend to cancel out one another.

The driving force behind the phase transformation is explained on the microscopic level by the decrease in free energy corresponding to the different interatomic bonds of the martensite as compared to the austenite below some critical temperature. Below this critical temperature, the probability that the interaction of the thermally vibrating atoms with a lattice phonon will give rise to a new configuration, corresponding to the martensitic arrangement of atoms, overcomes a threshold value [37]. The formation of the resulting distortion in the austenite occurs at velocities of the order of elastic wave propagation velocity in the material [38, 39], both in the longitudinal and the transverse directions, carried by the different vibration modes as they spread out from the newly formed martensite. Hence, the process is adiabatic [40], and as such the strain energy involved in the transformation is not relaxed in any significant manner, but rather stored in the material.

1.5 Physical and Mechanical Properties Featured in MPT and SME

Following is a list of definitions as adopted from [37]:

(a) Forward and reverse transformations: the austenite→martensite (or HTP→LTP) and the martensite→austenite (or LTP→HTP) transformations, respectively.

(b) Equilibrium temperature, T_E: the temperature at which the free energies of the austenite and the martensite of comparably high structural perfection are equal.

(c) Zero stress transformation, $\sigma_a=0$: transformation occurring in the absence of an applied stress.

(d) Martensite start temperature, M_S: the temperature at which martensite first forms on cooling the austenite in absence of stress.

(e) Martensite finish temperature, M_F: the temperature at which the transformation to martensite is completed on cooling at zero stress.

(f) Stress induced martensite start temperature, M_D: the highest temperature at which the transformation to martensite can take place on cooling under an externally applied stress.

(g) A_S, A_F, A_D: the temperatures corresponding, respectively, to M_S, M_F, M_D for the reverse transformation on heating.

(h) Hysteresis loop: in the present context, the strain-versus-temperature loop obtained on thermal cycling.

Aside from investigations connected to crystal structure, most of the research effort related to physical and mechanical properties of MPT and SME was concerned with the investigation and understanding of the different types of hysteresis loops [41]. Although a vast amount of experimental observations was accumulated during the past 20 years, some ambiguity exists when these results are compared. For example, Wasilewski defines M_S and A_S by intersection of slopes in a strain temperature (ε-T) hysteresis loop (Fig. 1, ref. [42]), while Sandrock et al. by first appearance of surface relief in an x-ray diffraction integrated intensity curve versus temperature (Fig. 11, ref. [43]). Further, although it is widely accepted that surface relief is one of the basic

fundamental properties of MPT [44], there is no prerequisite that the onset point of the different features involved in MPT be identical. As a matter of fact, it was pointed out by Wayman et al. that the resistivity vs. temperature (ρ-T) hysteresis loops are quite sensitive to prior history of the sample, and thus, should not be used for defining the A_S and A_F temperatures [45].

Even if there were a unique, widely established process to determine the characteristic temperatures, their values depend sensitively on alloying additions or impurities. The SME A_F temperature, for example, is found to be quite sensitive to minor substitutions of Co or Fe instead of Ni in Ni-Ti, causing a decrease in the recovery temperature. On the other hand, the substitution of Zr, Al, and Mn instead of Ti in Ni-Ti alloys, causes an increase in the A_F temperature [46-47]. The oxygen level affects the transition behavior [27] and also M_S [48-52].

Despite the above difficulties, general conclusions as to the characteristic temperatures of the phase transformation and the general shape of the hysteresis loops are available. As for the characteristic temperatures, rather than trying to define them from a single feature of the MPT or SME, the property most suitable to define them can be chosen. Thus, M_S can be defined as that temperature at which surface relief is observed [43], M_D, as the temperature "below which premonitory lattice oscillations are active and detectable" [53-54], M_F as "the highest temperature at which, when bending is applied to the sample, no further

spontaneous bending occurs on further cooling" [36], and T_E can be approximated as $T_E=(M_S+A_F)/2$ [55].

The width of the hysteresis loops seems to be affected by a number of mechanisms. On the crystallographic level, it is likely that the narrow temperature range is a consequence of small lattice shear on transformation [11]. Wasilewski attributes the ε-T hysteresis loop width in TiNi of 34°C to a relatively high interstitial content [56]. Using the concept of strain energy, accommodated by the austenite-martensite mixture on transformation. One can argue that the hysteresis loop width is a direct result of the different specific volumes of the austenite and martensite. Thus, as at a given temperature the fractional amounts of the martensite and austenite were different for forward and reverse transformations, the large hysteresis is observed [57].

1.6 The Stress Dependence of MPT and SME

1.6.1 Zero stress transformation

Under no applied stresses in a well-annealed sample with no preferred orientation, it is tempting to assume that all possible variants are formed with equal probability throughout the austenite in a self-accommodating manner. As the transformation is reversible, plastic deformation and hence creation of mobile dislocations are forbidden. The stored energy in the material that can play a role in the transformation is the elastic strain energy. It appears that the softening of the shear constant, $C'=(C_{11}-C_{12})/2$ corresponding to the resistance for the

$\{110\}\langle 110\rangle_{bcc}$ shear, as the sample temperature approaches M_S, has an influence on the lattice invariant distortions [58], as well as on the shear stress component necessary for the transformation [59]. At the same time, however, the new martensite product will continue to be affected by a restoring force, due to the existence of the elastic anisotropy, C_{44}/C' (Fig. 10 in ref. [58]). This restoring force gives rise to a frictional term in the total free energy change upon transformation [60-61].

The driving force for the transformation is the difference in free energies for the austenite and the martensite. The transformation will start when a critical supercooling is achieved. Once it starts, the free energy decreases progressively as the transformation proceeds. However, the transformation proceeds usually upon further decrease in the sample temperature [62].

As the transformation proceeds with an increase in the amount of martensite, a simultaneous increase of the internal stresses, σ_i, and thus of the unrelaxed stored elastic energy occurs. Since the free energy of the austenite-martensite mixture depends both on the temperature and on the shear stress during transformation, their effects are combined even in the absence of applied stresses [63].

1.6.2. Forward and reverse transformations following prestressed and prestrained conditions

It was delineated in the early stages of research on Ni-Ti alloys that prior plastic deformation plays a key role in the transformation

behavior. As Buehler et al. showed, an NiTi sample expanded or contracted on forward transformation according to the type of cold working it suffered prior to transformation. Swaged (compressed) alloys contracted, while drawn (tensed) alloys expanded [64]. Contraction on forward transformation was found by Wasilewski in hot swaged and annealed NiTi rod (annealed at 600°C for 1 hr), and he attributed this behavior to "the presence of significant texture and residual strains" [56]. Similar behavior was found for NiTi prestressed to 174 ksi in compression at 20°C [56].

The longitudinal dilation is not the only parameter affected by prior deformation. The simple peaks (one peak on forward, one on reverse transformation) of the internal friction and the elastic moduli became quite complicated in nature, following a 5% deformation of the NiTi sample at 80°C [65]. Hence, it is safe to state that the thermomechanical history of the sample affects to a large extent its transformation characteristics.

According to Gupta, the explanation for this thermomechanical history dependence for the reverse transformation can be understood on the basis of a model that assumes the coexistence of two variants [66]. In the martensite, the deformation results in the growth of the favorably oriented variant at the expense of the other, followed by the displacement of a coherent $\{111\}_M$ interface. Shape recovery then occurs upon heating.

Wasilewski experimented with compressive and tensile deformations applied with the sample in the HTP (300°C) and in the LTP [67]. From his figures (Figs. 1 and 2 in ref. [67]), it is clear that the amount of prior deformation (5%, 4.3% or 9.6%) does not affect critically the magnitude of the transformation strain (Fig. 2), but its effect is more pronounced when applied to the martensite (Fig. 2) than to the austenite (Fig. 1). Further, the direction of the transformation strain is opposite for the two cases. Following compression in the HTP, the sample elongates on forward transformation. However, following compression in the LTP, the sample contracts on forward transformation. Furthermore, upon prior tensile deformation, the transformation strains are reversed. The overall effect of prior deformation is summarized by Wasilewski as "introducing a finite residual stress opposite in sense to that which caused the deformation; this stress then affects any subsequent, nominally thermal, transformation to the other structure".

1.6.3. Stress applied to partially transformed material

The limited interrelation between applied stress and temperature in the process of martensite production is well-established [68]. When the stress is applied during the forward transformation, the nucleation of martensite from available variants will not be equally probable. Favorably oriented variants will be preferred [66]. Since the transformation under applied stress is reversible, the martensite production mechanisms are restricted to those that are inherent to MPT in the absence of applied stress, namely twinning (dislocations must be sessile, and considerable slip is not possible). Nevertheless, in certain alloys, the

martensite plates are found to be faulted, instead of internally twinned [69-71]. However, reversibility still exists, if the martensite plates retain the partial dislocations inherent in the faulted structures [54].

It has been confirmed experimentally that:

(a) The stress distribution in the sample affects the nucleation of martensite [11].

(b) Growth of favorably oriented martensite takes place under the influence of applied stress, achieved by the migration of the martensite-martensite interface ($\{111\}_M$ twin plane) and the martensite-retained austenite interface (habit plane) [36].

(c) The applied stress causes a reorientation of the martensite by means of favored twinning systems [36].

(d) Motion of twin boundaries in a given martensite variant is induced by the applied stress [36, 72, 73].

Perkins relates all the above effects to "elastic deformation induced transformation (EDIT) of the high temperature phase to martensite and/or of the existing athermal martensite" [54]. Thus, the applied stress will promote those variants of the transformation that most easily can be compensated by the applied stress for the shear distortion of the transformation. By the selection of preferred variants, work is done by the applied stress over the distance of the transformation, causing an equivalent reduction in stored energy [37]. Thus, the resolved shear stress, τ, contributes to the chemical force, (the derivative of the free

energy change per unit volume) weakening the resistance to accommodation of the transformation product [55].

1.6.4 Experimental methods used in the determination of the stress dependence of transformation behavior

The dependence of transformation behavior on stresses was investigated in the past by the use of different experimental methods. These methods can be classified according to the kind of stress applied to the material and according to the existing phase in the material at the time of application of the stress. The experimental results of investigations using these methods should be compared with results for originally stress-free materials. In such materials, residual stresses are removed by annealing, followed by thermal cycling under zero applied stress [43].

Experiments of another type were done on prestrained material. The prestrain deformation can be induced either by tensile stresses, as in drawing [64, 67] or by compressive stresses, as in swaging or rolling [64, 67]. The plastic deformation process can take place in the HTP [67], in the fully martensitic LTP [67, 74, 75], or at an intermediate stage of the phase transformation [64]. In all the above cases, the induced residual stresses are aligned more or less with one of the sample axes. However, bending introduces both tensile and compressive stresses. As the effect of residual stresses depends on the stress distribution as well as on the phases present, thermal cycled samples that suffered bending at various stages of their transformation [36, 76, 77] exhibit different strain behavior from those that were swaged or drawn.

Stress can be applied dynamically [78] or statically, using dead weight [78-80] prior to transformation cycling and retained during it. Dynamic stressing can be carried at constant temperature and strain rate [58] (as for the usual stress-strain curve) or at constant strain during thermal cycling [81]. Dynamic stresses can be introduced also an an unavoidable consequence of an experimental procedure, as in internal friction measurements [65].

Only in rare cases was dead-weight loading used in the determination of the dependence of transformation strain on applied stress. In one experiment, a low constant load was applied to an Ni-Ti spring in the HTP, and then the load was increased after transformation to the LTP [59]. Thus, the dead-weight load was not kept constant for the complete thermal cycle. Complete dead-weight cycling was conducted, however, for Fe-Ni-C alloys [78] and for Ni-Ti tubes [79] and wires [80]. The Ni-Ti tubes were swaged from arc-melted and cast material, and the Ni-Ti wire was drawn from arc-melted and cast stock and then transformation fatigue cycled for various amounts up to millions of cycles. The tubes and wires were stressed only in tension, and both forms exhibited an increase in transformation strain with increasing applied stress. In the present experiments, the samples were of tubular and square cross-sectional shape and fabricated from powder metallurgical stock. Also, axial and transverse strains were measured upon thermal cycling as a function of tensile and compressive applied stresses.

2. EXPERIMENTAL PROCEDURE

2.1 General

Simultaneous measurements were made of the change in axial and circumferential (transverse) strains upon phase transformation, induced by temperature cycling under uniaxially applied load on tension and compression of Ni-Ti (~55 wt% Ni) or Nitinol material. The mechanical response of tubular and square cross-sectional solid bar material at low rates of change of temperature, caused by the interaction of residual and phase transformation stresses (both of them are internal to the sample) with the applied uniaxial stress, was investigated.

2.2 Materials and Equipment

2.2.1 The Nitinol material

The experiments were carried out on two tubular samples, PO-4 and P8-1, and the square cross-sectional solid bar, P7-1, fabricated at Ames Laboratories, (AL), from stock bars designated as #5, #8, and #7, respectively, provided by the Special Metals Company (SM). The final dimensions of the fabricated tubular samples PO-4 and P8-1 are shown schematically in Figs. 2(a) and 2(b), respectively. A photograph with a metric scale included shows sample P8-1 (Fig. 3), where the shoulder sections are clearly visible. The shoulder section is designed to permit the attachment of grips and to define the gage length. The gage length portion of the sample was centerless ground on the outside to give a

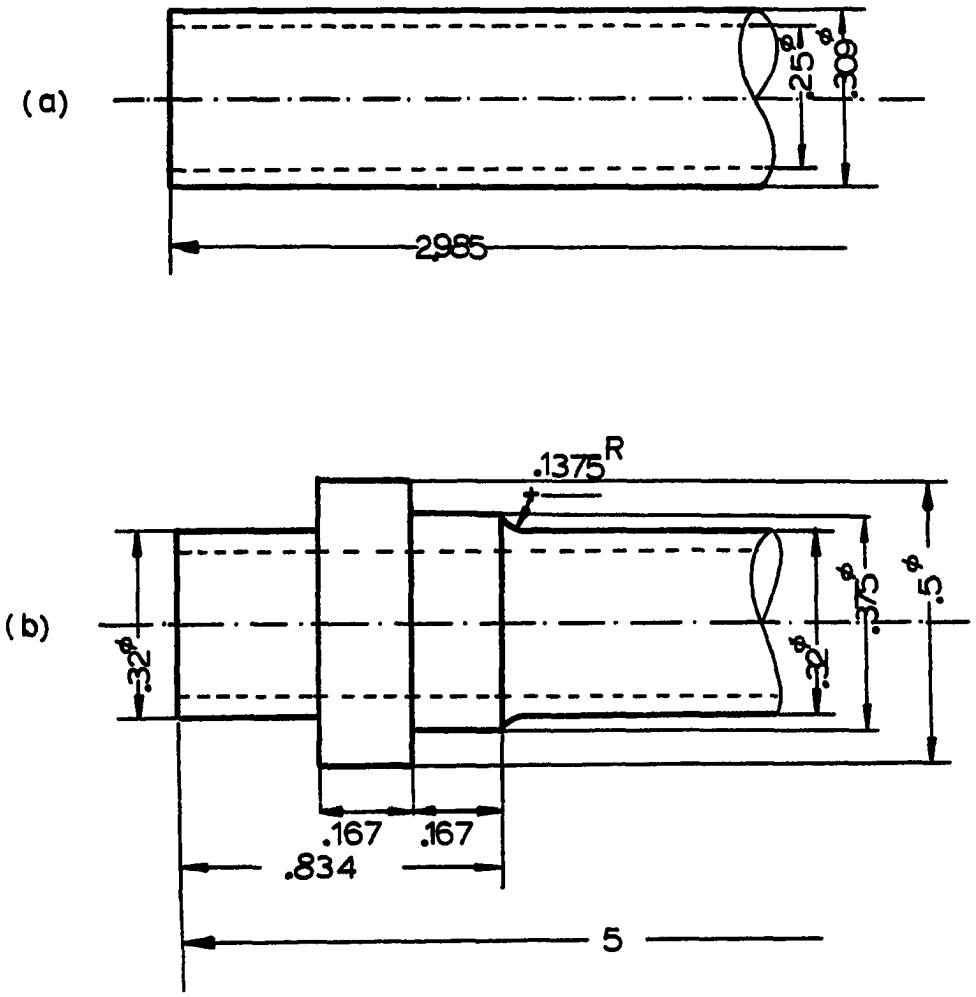


Figure 2. Schematic drawing of tube P0-4(a) and tube P8-1(b). All dimensions are in inches.

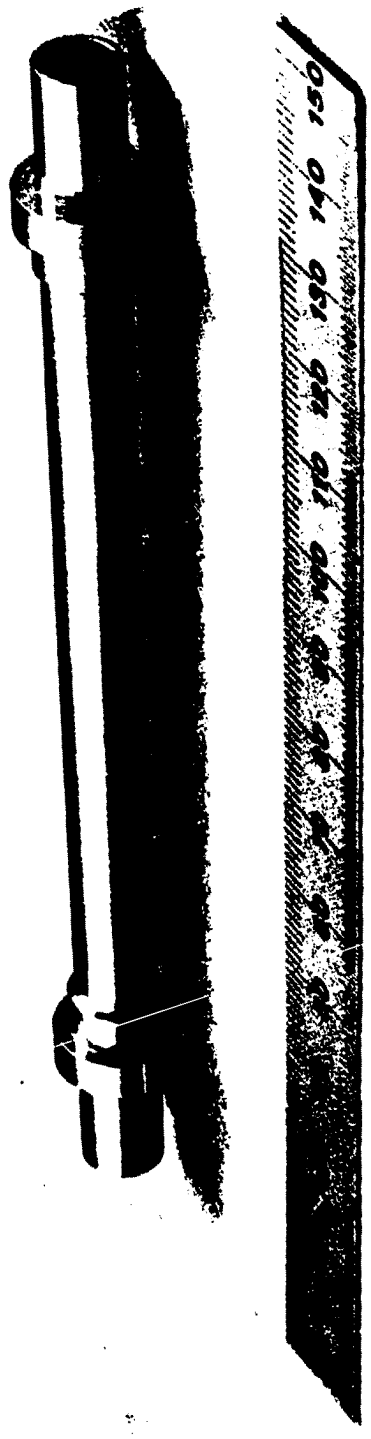


Figure 3. Photograph of sample P8-1 in the as-fabricated state

smooth surface. Fig. 4 shows the square cross-sectional bar, P7-1, of 0.265 x 0.265 inch, mounted with strain gages and a thermocouple.

The SM process for preparing the bars included the following steps [82]. Atomized Ni-Ti powder, with a predetermined Ni content, was canned and pressed by hot isostatic pressing (HIP). The resulting sintered bar was heat treated at about 900°C, followed by furnace cooling. The canning material was partially removed. In this condition, the bars were delivered to AL for further processing and investigations.

2.2.2. Chemical analysis

Several analytical methods were used to determine the atomic fraction of Ni present in the alloy (which mainly governs the phase transformation temperature) and to identify the major trace elements (which may happen to be the contributors to secondary phase precipitates). For each method, one or more samples were provided, taken from different positions along the axial direction of the stock bar of interest. A list of the methods used and the analytical values are given in Table 1 for the case of sample PO-4. It was found that bar #5 contained 54.58 wt% Ni and 45.43 wt% Ti, while bar #8 contained of 53.34 wt% Ni and 43.79 wt% Ti, with 2.87 wt% of unidentified elements.

2.2.3. Microstructure

An attempt was made to characterize the Nitinol material microstructure in the as-received, cold swaged, and heat treated conditions, using both light and scanning electron microscopy. The sample preparation procedure included: hot mounting in plastic or copper, grinding on silicon

Figure 4. Photograph of sample P7-1, mounted with strain gages and thermocouple

Table 1. Chemical composition of SM stock bar #5 (used for sample PO-4). All analyses were conducted at AL

Element Chemical Symbol & Units of Concentration	Ti [wt %]	Ni [wt %]	Co [ppm]	Al Ba Ca Cr Cu Mg Mn Si Zn	C [ppm]	O [ppm]	H [ppm]	N [ppm]
Amount of Element Found	45.43	54.58	500	Traces	666	1100	39	47
Weight of the Sample that was Used [gm]	2.030				3.637		0.982	
Method of Determination	Wet Chemical Analysis		X-Ray Emission Spectrography		Combus-tion	VF[a]	VF	VF

[a]Vacuum Fusion.

carbide grinding paper up to 600 grit, and polishing mechanically through Linde B (0.03 μm alumina) powder. In most cases, the samples were etched, using either 9.8 volume percent HNO_3 + 1.9 volume percent HF (slightly etched samples) or 5 volume percent HF aqueous solution under constant agitation in ultrasonic bath (heavily etched samples).

In the case of rod #5 (from which the tube PO-4 was prepared), the microstructure showed fairly equiaxial circular grains (Fig. 5(b)). A wide range of grain sizes was present, from grains with characteristic diameter as small as 2 μm and up to about 60 μm for samples. Each grain boundary contains large numbers of acicular, elongated entities of submicron size. These particles, which are too small for EDAX (Energy Dispersion X-Ray Analysis) identification of their composition, appear on the photomicrographs as bright inclusions in the matrix material at magnification of 1000X (Fig. 6(c)). Apparently, the same particle also forms a delicate, interrupted, discontinuous network inside the grains, in such a manner as to subdivide the bigger grains into domains of nearly equal size of about 6-8 μm. Seemingly, in each domain the martensitic plates are unidirectional, but their direction is changing vastly from domain to domain in the same grain. The subdivision of smaller grains appears to be finer, while the smallest grains do not seem to be subdivided at all, although they do contain one or two precipitate particles (Figs. 6(c) and 6(d)). The above-mentioned features are observed for material cold swaged from the stock bar with reduction of area (RA) of about 11% (Fig. 6) and also for the as-received material following

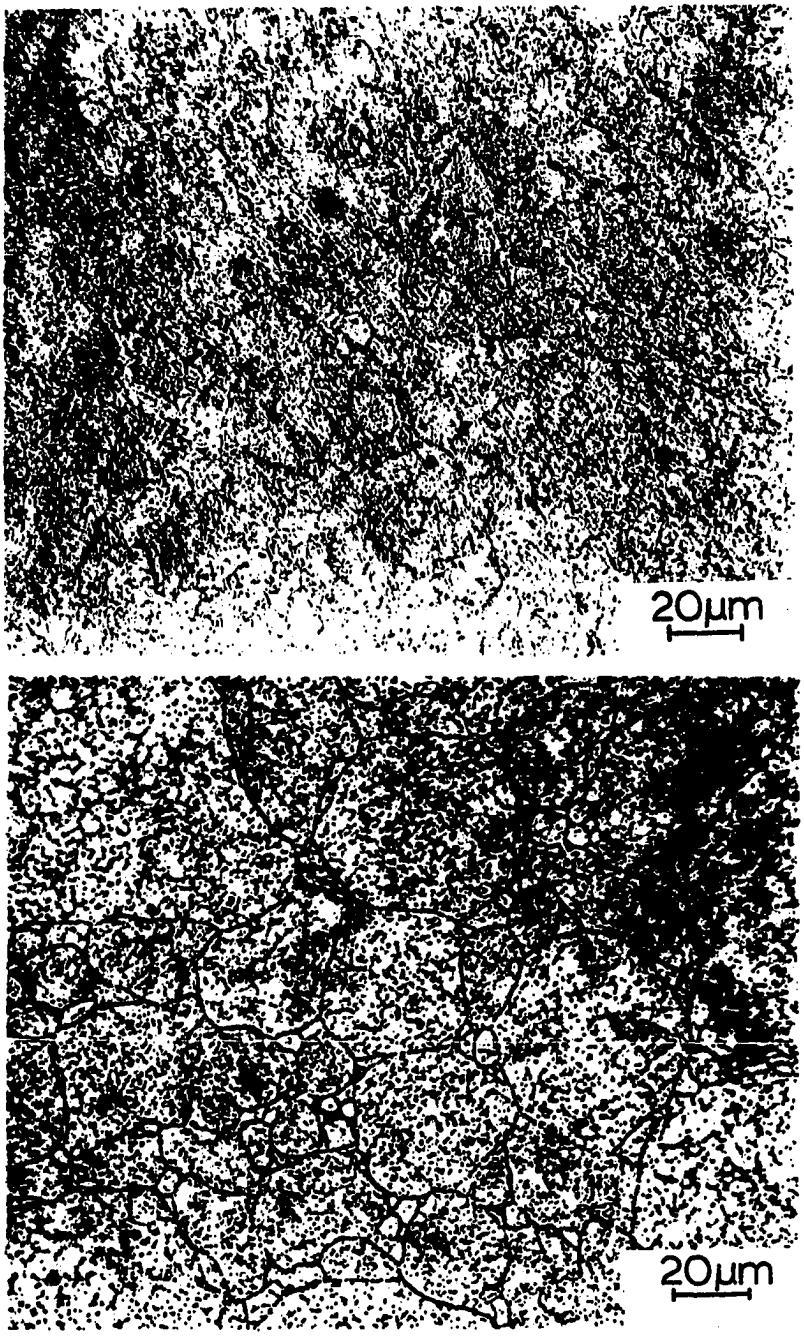


Figure 5. Lightly etched (9.8 vol% HNO_3, 1.9 vol% HF) transverse section of as-received stock bar # 5 material (a), and the heavily etched (5 vol% HF) counterpart(b) showing equiaxial grains of varying grain size and secondary particles

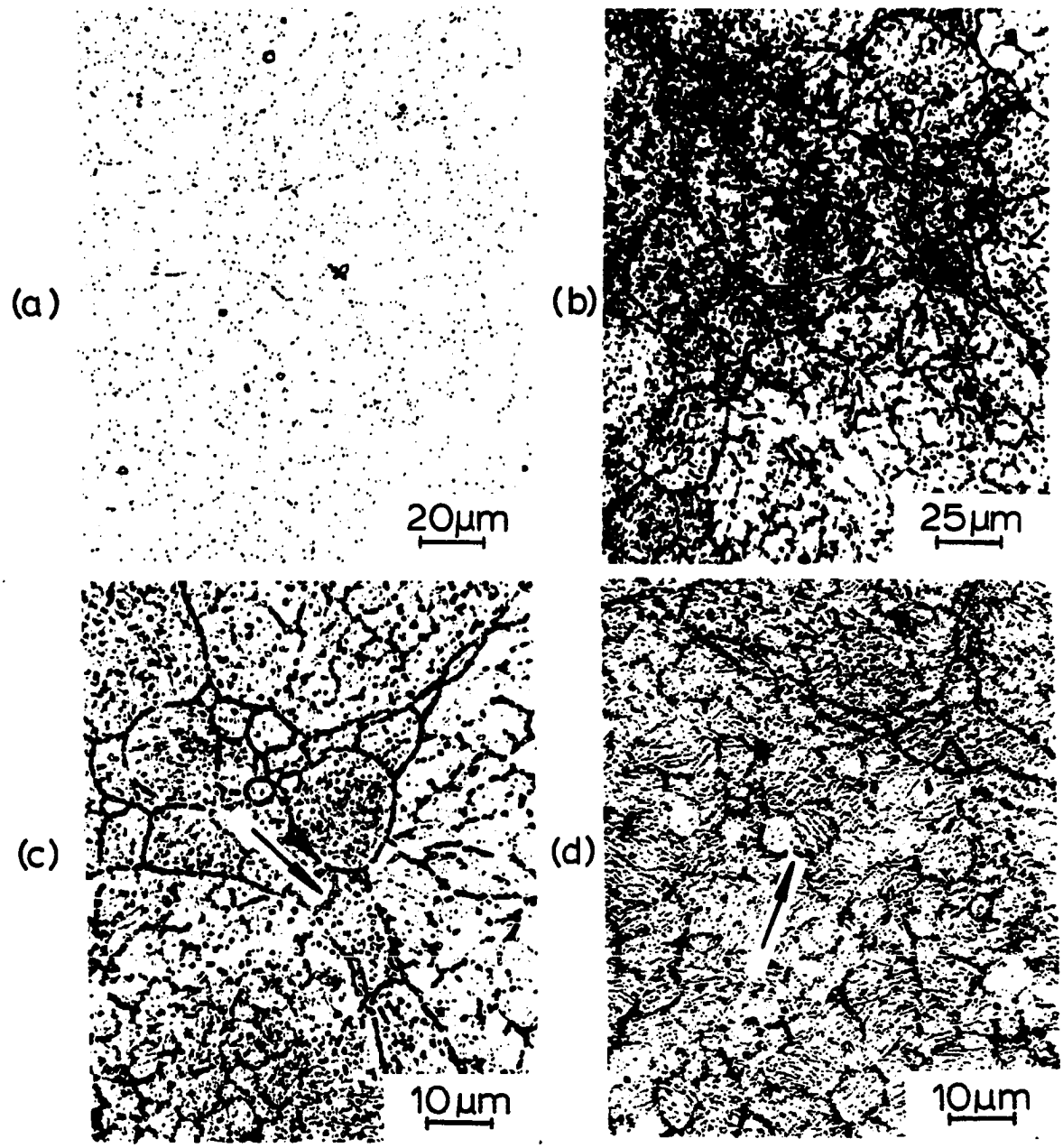


Figure 6. Photomicrographs of transverse sections of a cold swaged (11% reduction of area) tube from stock bar #5. The lightly etched surface, 9.8 vol% HNO_3+1.92 vol% HF, (a), shows precipitated secondary particles which form "domain" boundaries. The heavily etched surface (5 vol% HF) indicates the grain boundaries (b). At higher magnification the precipitated network structure is revealed (c) together with "domain" formation caused by well oriented martensite plates (d)

annealing in vacuum at temperature of 600°C for 4 hours and furnace cooling prior to the sample preparation procedure (Fig. 7(d)).

Another particle was occasionally observed (see arrow, Fig. 8). These polyhedral, mostly rectangular, bright features are much bigger than those forming the network. Their size is about 1-3 μm, and is independent of the grain size. They may have been formed by coalescence and growth of the previously mentioned much finer precipitates, and maybe they are of totally different origin and composition (Fig. 8). Their size makes it possible to analyze them by EDAX, the result of which is shown in Fig. 9. According to this result, the particles in question are rich in Ti relative to the surrounding matrix, giving rise to the assumption that they may be intermetallic particles of composition Ti_4Ni_2O [26].

Fractography of transverse section upon cold-drawing of a previously cold-swaged and heat treated sample shows alternating ridges, which are oriented 45° to the crack propagation direction (Fig. 10(a)). Although equiaxed dimples predominate (Figs. 10(c) and 10(d)), some indication for local cleavage was also found (Fig. 10(b)).

The nickel to titanium weight ratio of stock bar #8 was higher than that of stock bar #5 (1.22 compared to 1.20), and therefore M_S for stock bar #8 material was lower than for stock bar #5 material. As a result, a martensitic structure was rarely observed under the microscope at room temperature, either in the as polished or in the etched condi-

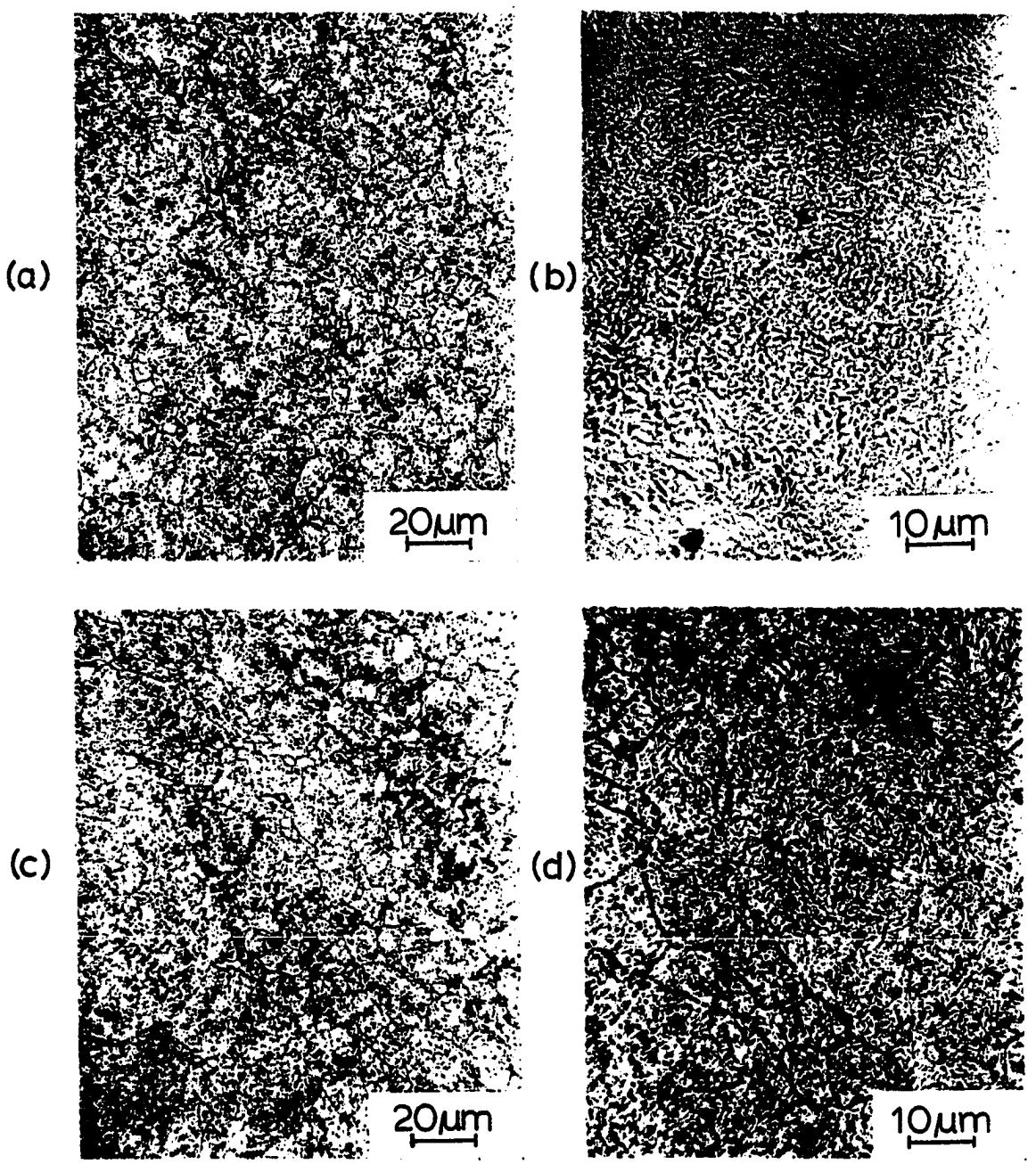


Figure 7. Photomicrographs of two transverse sections sliced at different axial positions from stock bar #5 in the as-received condition ((a) and (c)), and after heat treatment in vacuum at 600°C for 4 hours followed by furnace cooling ((b) and (d)), polished, and heavily etched, respectively. The martensitic plates are clearly observable at the higher magnification for both the etched (d) and the polished (b) conditions

Figure 8. Heavily etched (5 vol% HF) transverse section of cold-swaged (RA = 11%) Nitinol showing some bright, angular, secondary particles in the middle section of the micrograph

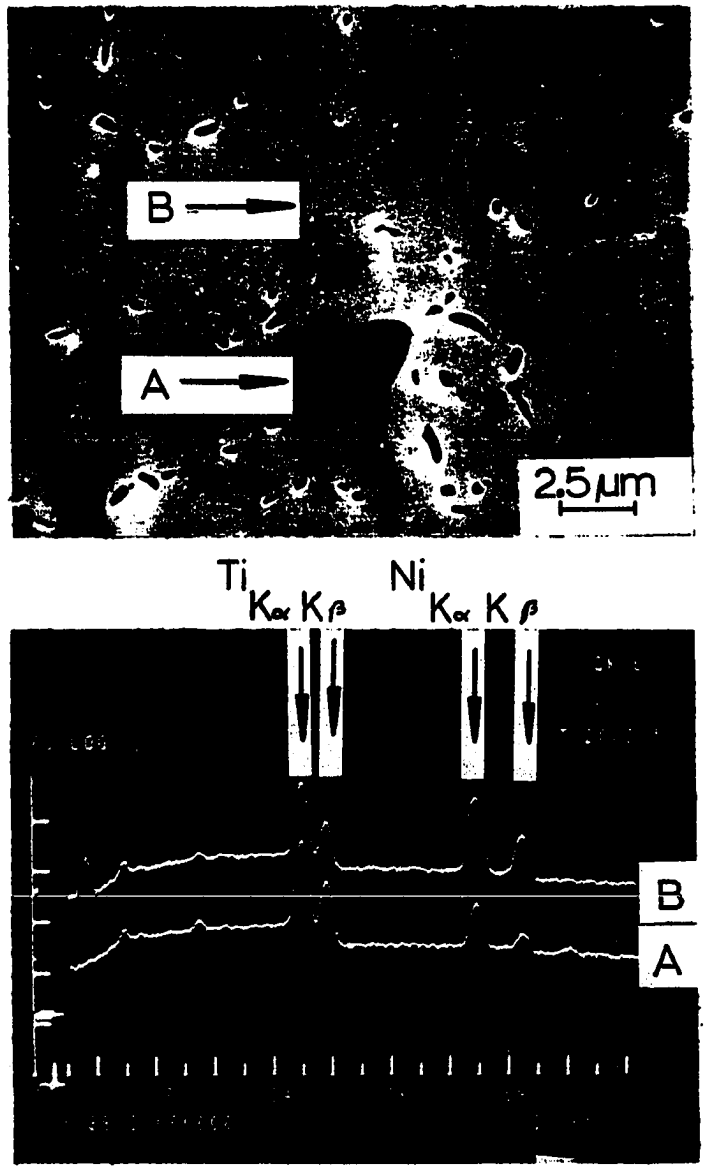


Figure 9. Selected area SEM photomicrograph (a), showing secondary particle A and matrix material B. The secondary particle was identified by EDAX as rich in Ti compared to the matrix (b)

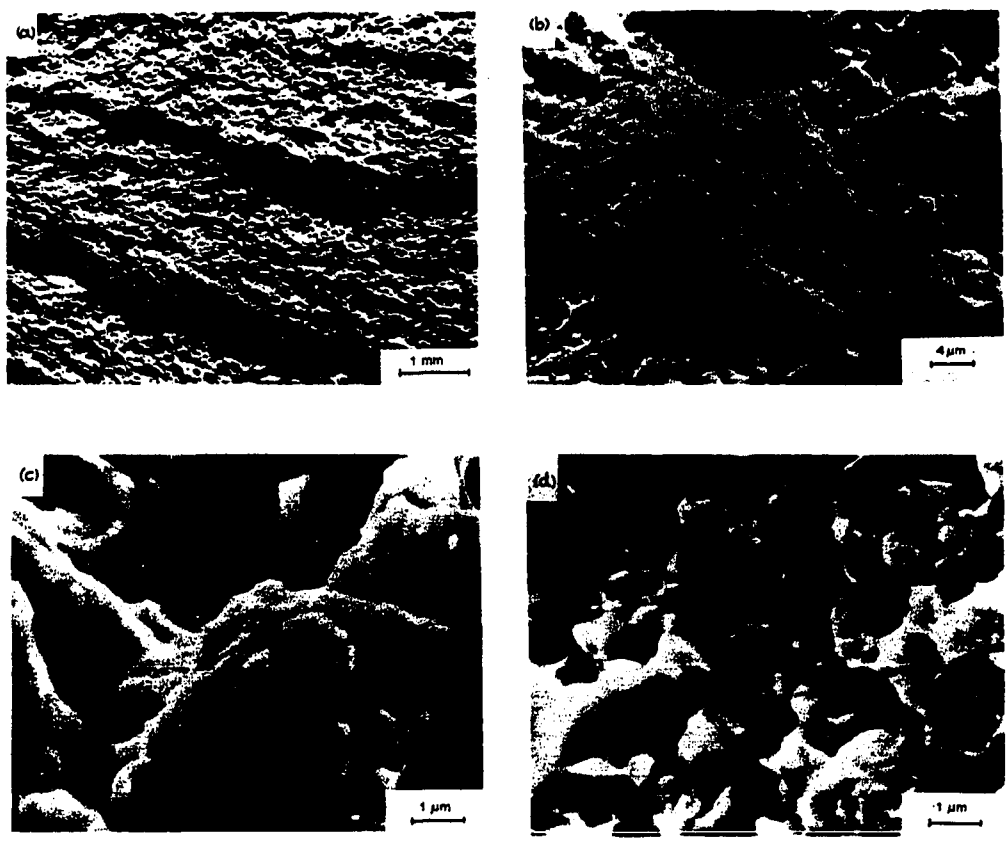


Figure 10. SEM fractomicrograph of a transversely sectioned sample from a cold-swaged, annealed, and cold-drawn tube (upon which cracking occurred) at a transverse crack near the tube outer surface. Alternating ridges (a), local cleavage surrounded by dimples (b), and the morphology of the dimples ((c) and (d)) are shown

tions. However, around microscopic imperfections, martensite formation did occur, presumably stress induced, as shown in Fig. 11. The microstructure was found to be very similar to that of stock bar #5 with respect to grain size distribution, average grain size, and the equiaxial character of the grains. As in the case of stock bar #5, no apparent differences were found between microstructures taken along the bar axis and in the perpendicular direction. In Fig. 12, the etched microstructure of stock bar #8 is shown at two magnifications, for the longitudinal section ((b) and (d)), and for the transverse section ((a) and (c)). The major differences between the two microstructures, that of bar #5 and bar #8, consist of two features:

a) A larger number of secondary particles in bar #8, mostly of acicular shape (Fig. 12(c) and (d)), resolvable both for the as polished, and polished and etched conditions (Fig. 12).

b) What seem to be closed pores in bar #8, mostly along grain boundaries.

2.3 Heat Transfer Systems

2.3.1. The oil bath

Sample P7-1 was investigated under conditions of immersion in a heat reservoir. The oil temperature bath was made of a thin-wall stainless steel cylinder with fluid content of about 5 gallons. The heat transfer liquid was Dow Corning 200 silicone oil. The bath was heated by an elec-

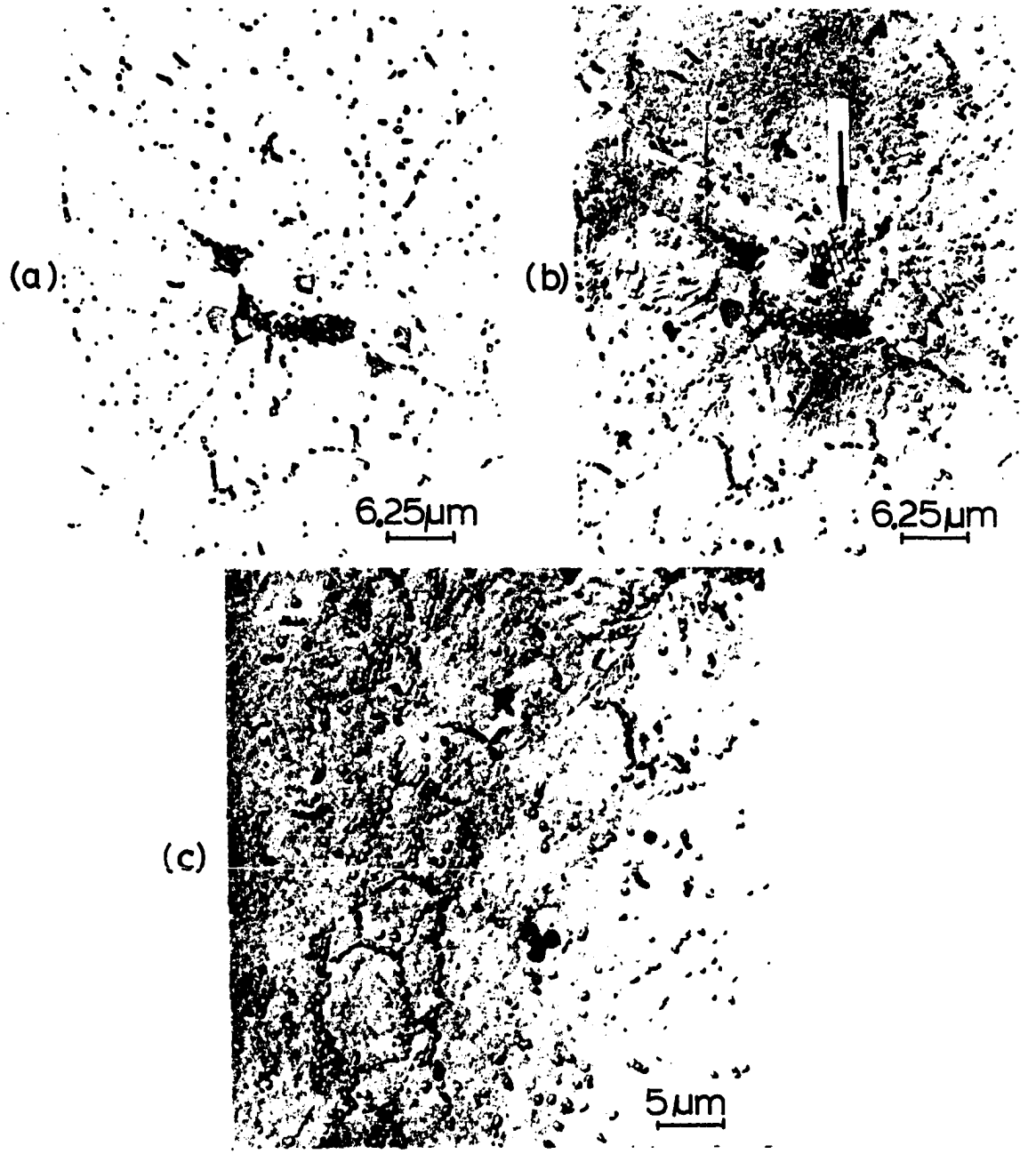


Figure 11. Heavily etched (5 vol% HF) transverse section of stock bar #8. (a) Bright field photomicrographs showing different kinds of inclusions and related presumable stress induced martensite near them. (b) Differential interference contrast of field shown in (a). (c) Enlarged center of (b)

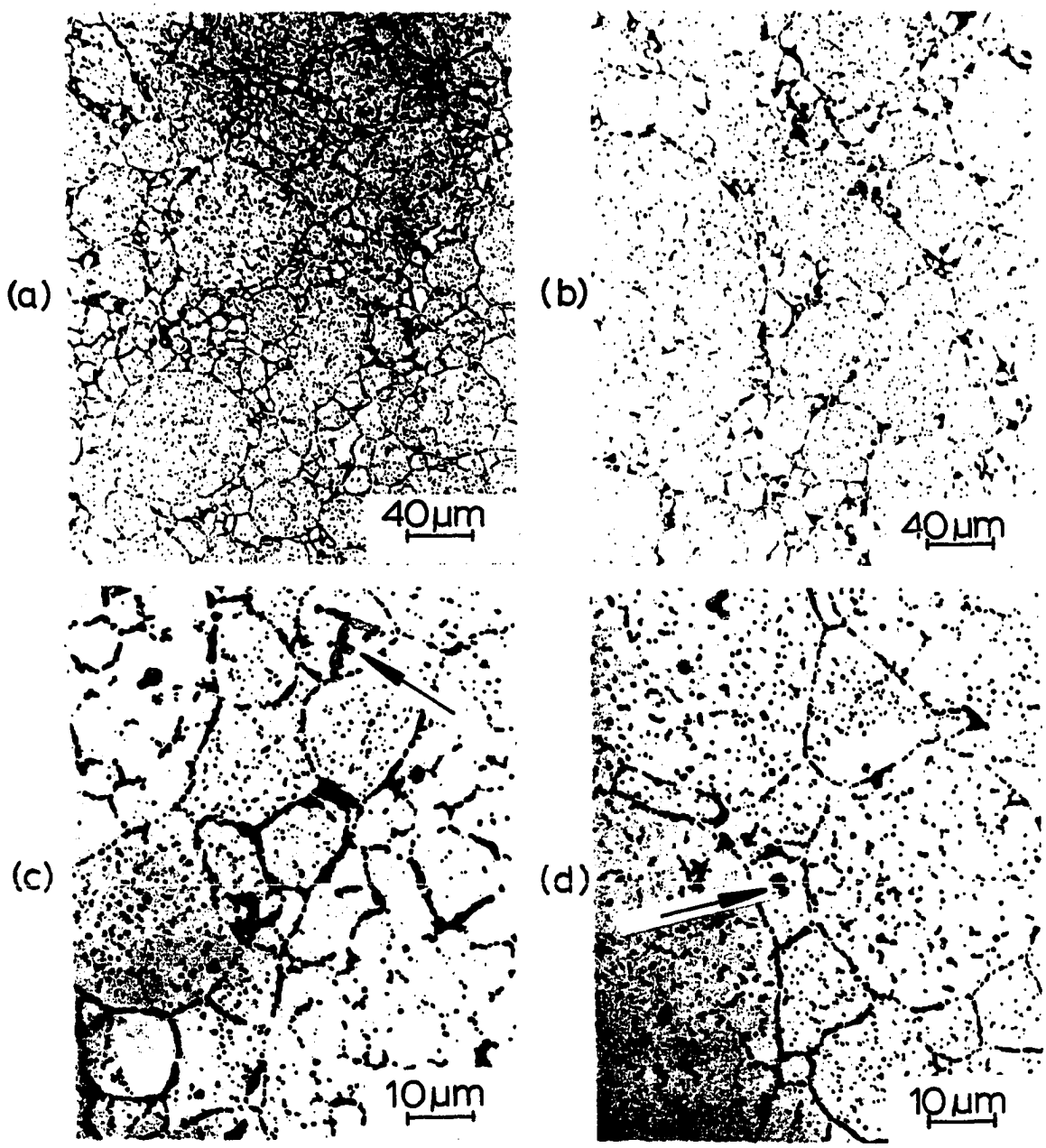


Figure 12. Photomicrographs of heavily etched (5 vol% HF) surfaces, along the bar axis ((b)) and (d)) and perpendicular to this axis ((a) and (c)), from stock bar #8. No differences can be observed between these two sections. At the higher magnification of 1000X high density secondary particles are resolvable ((c) and (d)). The microstructure is equiaxial. Wide range of grain size distribution is exhibited

tric immersion heater and cooled by a cooling coil through which water or cold nitrogen gas was pumped.

2.3.2. Constant temperature fluid recirculation unit

A standard MGW-Lauda type NS-HT unit was used, with maximum available heating power of 2070 W. The flow capacity of the system with the Dow Corning 200 silicon oil was about 3-5 liters/min, which means an average residence time of 6×10^{-2} sec of the heat transfer fluid in the tubular samples. The fluid was pumped from the reservoir, through the sample and the connecting tubes, and back to the reservoir in a closed circuit.

2.3.3. Temperature control

In the case of the standard MGW-Lauda recirculation unit, temperature control was achieved by balancing the heating power and the fluid volume in the tank on heating and the cooling water flow rate and the driving pressure in the liquid nitrogen cryostat flask on cooling. During the heating part of the thermal cycle, the temperature of the fluid in the reservoir was raised using an immersion heater, which raised the sample temperature at a rate of 5°C/min. For cooling down to room temperature, water was circulated through a copper coil immersed in the fluid. For cooling below room temperature, cold nitrogen gas was circulated through the cooling coil. The cooling rate was more temperature dependent than the heating rate, with an average of approximately 2°C/min.

The oil bath temperature was controlled on heating by an MGW-Lauda contact type, thin glass, mercury thermometer immersed in the oil bath.

Its movable electrical contact control was attached to the shaft of a stepping motor. The contact leads themselves were connected to a Precision Scientific electrical relay. This relay, upon command from the thermometer, opened and closed the power circuit to the electrical heater. A schematic diagram of the electrical connections is shown in Fig. 13.

2.4. Thermocouples

All the thermocouples consisted of Omega Engineering G-G-T-24, glass-braid insulated, T type (copper-constantan) thermocouple wire. They were connected through a cold junction solid-state compensator to an Omega 2809 digital display unit, for direct survey of the temperatures of interest.

One of the sensing thermocouples, firmly spot welded to the surface of the sample near the strain gages and thermally insulated from its outer surrounding, was wired through a Vishay 1611 thermocouple adapter to the Vishay Instrument Division 220 data logging system.

2.5 Strain Gages and Data Logging System

Strain gages, made by Micro-Measurements of the rosette CEA-06-062UR-120 (tubular samples) or CEA-06-062UW-120 (on sample P7-1) types, were used for sensing changes in strain upon phase transformation of the sample material. In Fig. 14, a schematic diagram of the rosette type strain gage is shown mounted on a section of the tubular sample. Temperature calibration curves are provided with the strain gages, as deter-

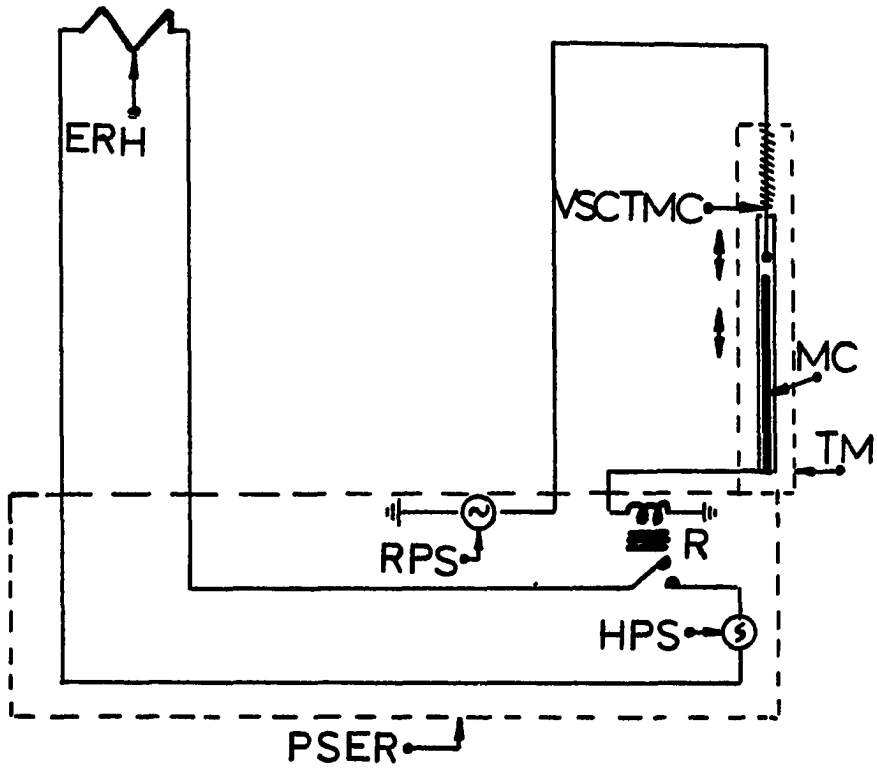


Figure 13. Schematic layout diagram of electrical connections of temperature control of oil bath upon heating showing variable speed controlled thermometer contact (VSCTMC), expandably mercury column (MC), Lauda contact type thermometer (TM), electric resistance heater (ERH), power supply to relay and heater (RPS and HPS, respectively), and the Precision Scientific electric relay case (PSER)

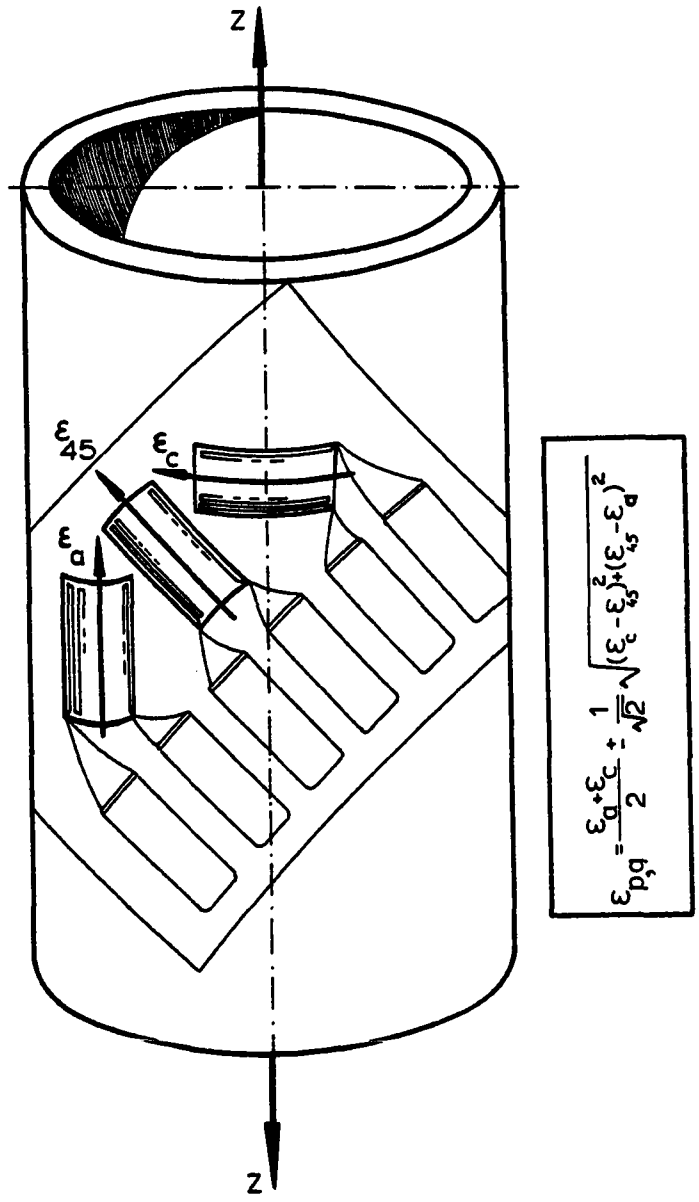


Figure 14. Schematic of rosette type-strain gage mounting on tubular sample section. The equation gives the relationship between major and minor principal planar strains, ε_p and ε_q, respectively, and the measured quantities, ε_a, ε_c, and ε_{45}

mined using 1018 steel as the substrate material. Since the thermal expansion coefficient for 1018 steel, $11.8 \times 10^{-6}(°C)^{-1}$ [83, 84], is close to that for HTP equiatomic Ni-Ti, $1.1 \times 10^{-6}(°C)^{-1}$ [27], the provided calibration curves should be valid here. The calibration curve indicates a maximum temperature-dependent change in apparent strain of less than 1.1×10^{-4} or 0.011% for the entire temperature range from about -30°C to 150°C. Therefore, no temperature corrections to the strain gage outputs were applied. The strain gages were bonded to the surface of the electropolished sample near its midsection. Their axes were aligned along the axis of the sample and along its transverse direction (perpendicular to the symmetric axis of the sample). The strain gages were selected from the same package in order to cancel out errors due to manufacturing parameters. The strain gages were cemented to the sample surface by epoxy resin, cured at about 175°C for 1 hour under vacuum and furnace cooled. Wires were soft-soldered to the strain gage contacts, and the whole assembly was covered by a flexible moisture-and-electrical insulating self-curing resin. The leads of the strain gages and the thermocouple were wired to the Vishay Instrument Division 220 data logging system. Connections were made in the quarter-bridge mode to enable simultaneous and individual measurements of the axial, circumferential, and the 45°-inclined strains. Phase transformation was induced by thermal cycling under uniaxial load, and the strains were measured as a function of the sample temperature. The schematic connections to the V/E-20A

strain indicator (which is part of the data logging system) are shown in Fig. 15 [85].

2.6. Auxiliary Equipment

An effort was made in the design and construction of the supporting device to avoid undesirable bending moments during thermal cycling. The supporting grips had to transfer loads up to 350 pounds in a uniaxial manner and at the same time maintain a tight seal on the hydraulic connections. The shoulder sections on the sample (Figs. 3 and 4) with one surface acting as load-carrying, while the other as sealing, served the above purposes.

The sample equipped with the sensors for strain and temperature measurements was inserted into two pairs of split grips, one grip at each end of the sample. Premachined grooves in the grips seat against the shoulders on the tube to maintain load carrying ability, while flexible tubing attachments to both ends of the sample through windows in the split grips provided the leak tight circulating fluid connections. With the sample in a vertical position, one pair of split grips at one end was connected through a universal joint and shearing pin to a load-carrying heavy steel frame, while the other pair, with an identical arrangement, was connected to the loading tray, which carried the dead load in the form of preweighed lead bricks. The schematic diagram of one pair of the installed split grips is shown in Fig. 16.

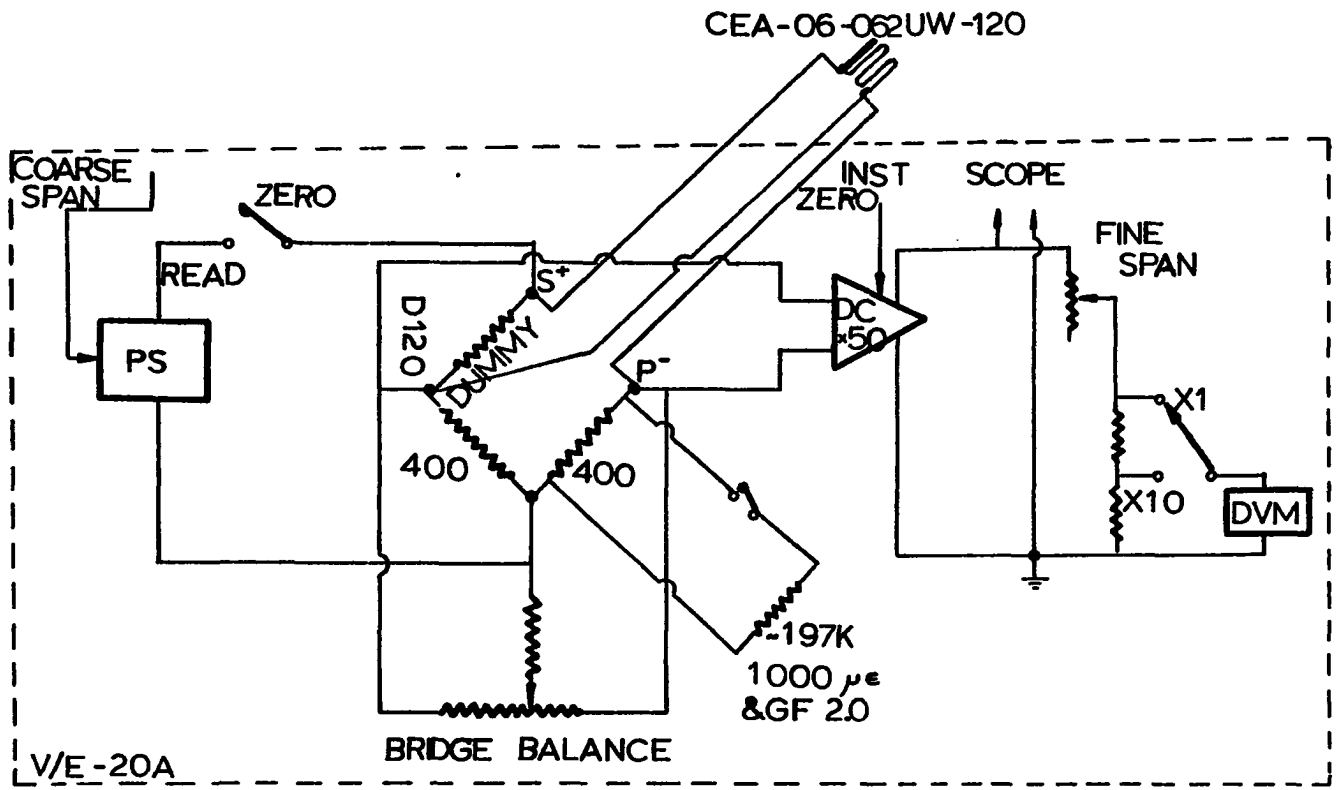


Figure 15. Schematic electronic layout diagram of V/E-20A strain indicator with a CEA-06-062UW-120 type strain gage connected to it in the quarter bridge mode [85]

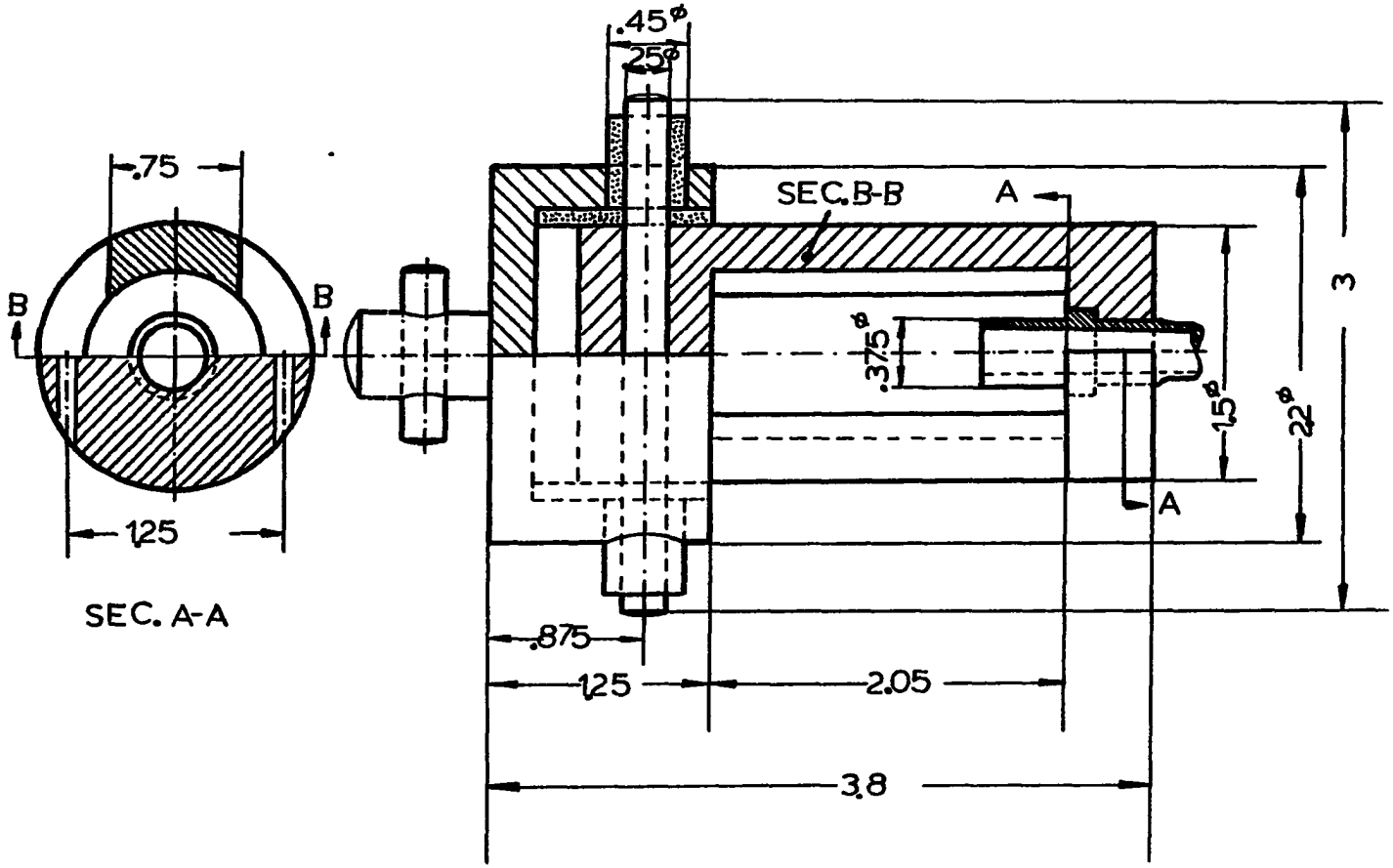


Figure 16. Suspension grips for varying deadweight loading experiments. The sample is electrically insulated from the universal joints(not shown). Only major dimensions are given (in inches) for sake of clarity.

Some of the experiments were performed in compression, and so a different type of grips was used. A pair of split grips was mounted vertically on a flat surface of a heavy steel base through a tubular, rotating extension, while the other pair was able to glide along two parallel vertical columns by means of roller bearings. Thus, the sample movement was constrained in all directions, except along a single vertical line through its axis. The load-carrying tray was firmly attached to the gliding grip. The compression device and the lower split grip are shown in Fig. 17.

2.7 Experimental Procedure

2.7.1 Calibration of measuring devices

The display reading (number of counts) of the V/E-20A strain indicator was calibrated versus magnitude of input voltage signal. The calibration was carried out by sending currents up to 100 mA from a constant current source through a standard resistor of 1Ω. The voltage drop across the resistor was registered by a Keithley-180 nanovoltmeter. The calibration curve is shown in Fig. 18.

The temperature-time behavior of sample P8-1 during thermal cycling is given in Fig. 19. It was found that the average rate of heating between -30°C and 85°C was 5.12°C/min, and on cooling between 25°C and -30°C it was -1.66°C/min, using a least-squares approximation. The cooling rate between 85°C and 25°C was approximated with a parabolic function as given by equation (3),

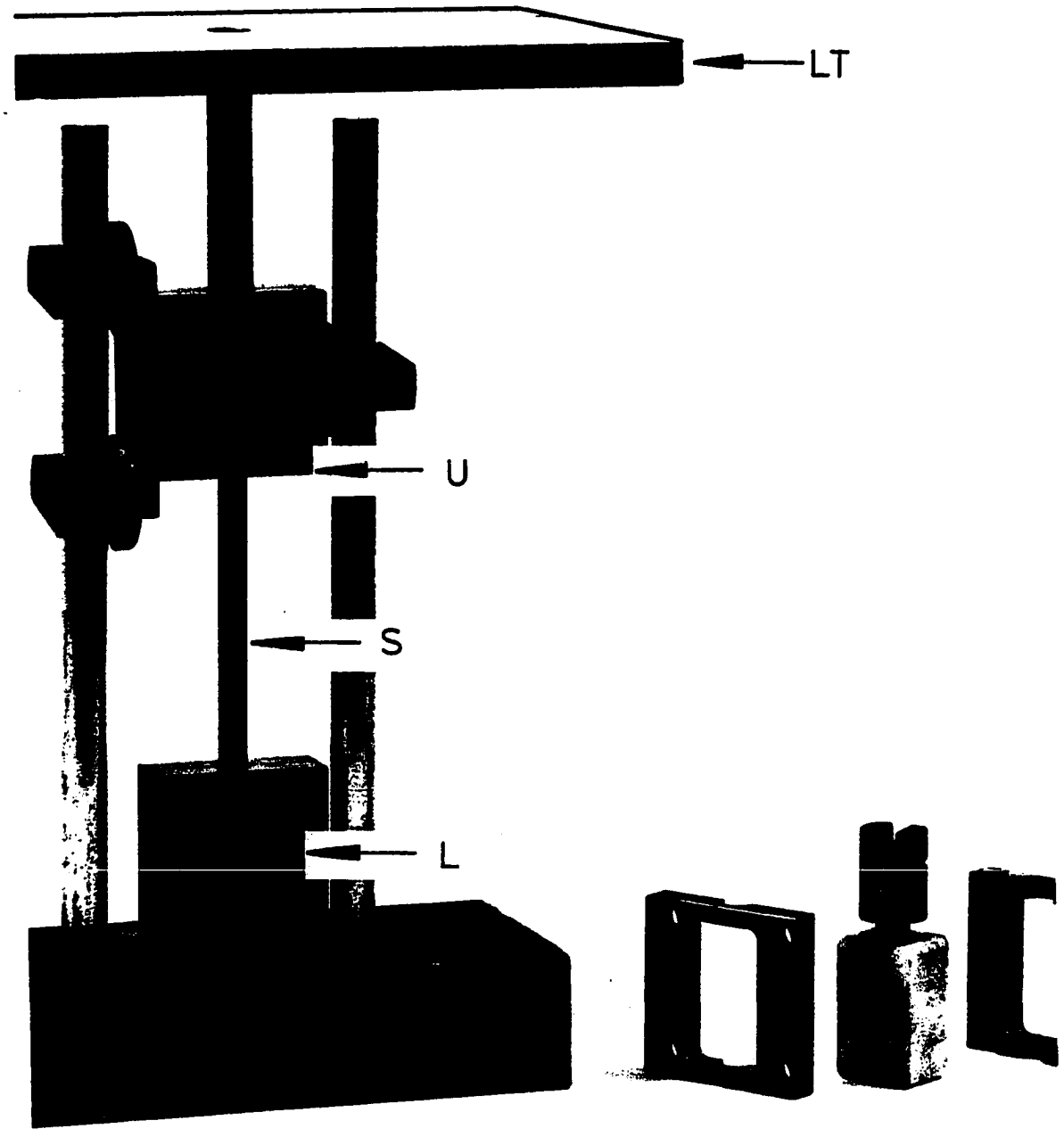


Figure 17. Compression device photograph, showing lower fixed grip (L) and upper gliding grip (U) to which the loading tray (LT) is attached. Thus, the sample (S, between the two grips) is able to move freely only along a single vertical line. To the right, the splitted lower grip is shown

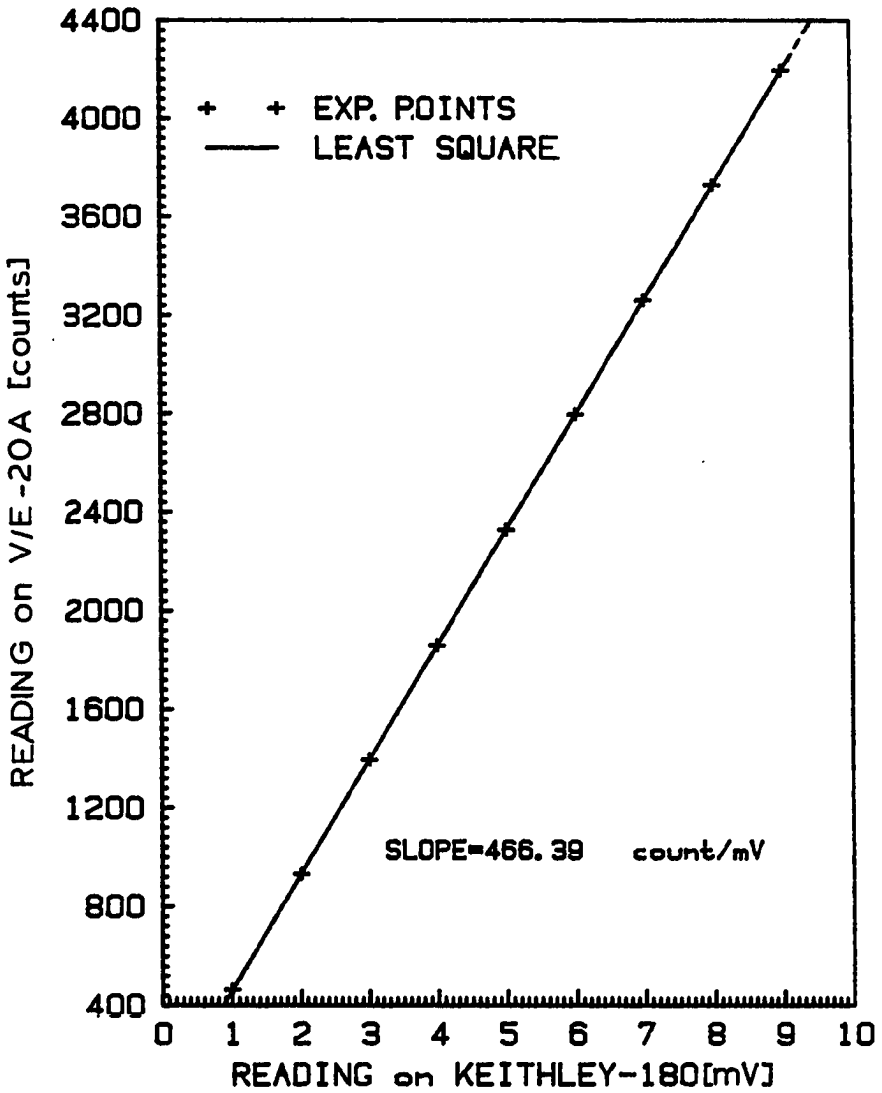


Figure 18. Calibration curve for V/E-20A strain indicator

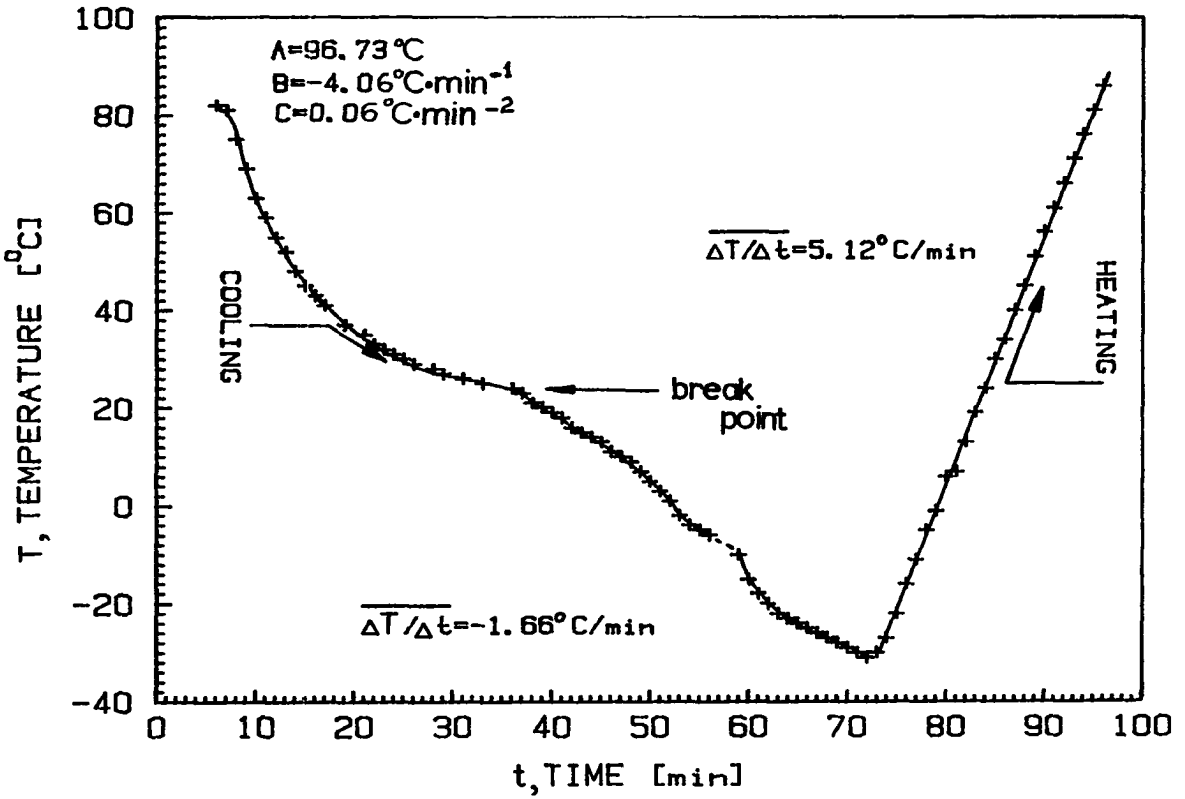


Figure 19. Typical temperature-time profile of sample P8-1 outside surface on heating and cooling, with average heating rate of 5.12°C/min and cooling rate of 1.66°C/min (during nitrogen vapor circulation part of the thermal cycling)

$$T = A + Bt + Ct^2 \qquad (3)$$

where T = temperature in °C

t = time in minutes

and the constants have the values:

A = 96.73°C

B = -4.06°C·min^{-1}

C = 0.056°C·min^{-2}

2.7.2 Experimental sequence and data taking procedure

The mechanical response of sample P8-1 was investigated first in the as-fabricated condition followed by the curing thermal cycle for the strain gage cement. In this condition, simultaneous measurements were made of the change in axial and circumferential strains as a function of temperature upon phase transformation under various uniaxial loads. The runs were started at zero load at room temperature, where the strain gage output was zeroed, followed by heating the sample at a nearly steady rate toward a target limit where the circumferential strain gage output showed a linear increase with increasing temperature. It was assumed that at this point the sample had transformed to the high temperature B2 phase. The circumferential strain gage output was preferred over the axial one for setting the temperature limits because of its higher fractional sensitivity than that of the axial strain gage. If the run was to be under a uniaxial applied stress, then the temperature was stabilized within ±1°C near the high temperature limit (85°C), and the loading tray was

attached to the lower grips. The tray was then loaded, gradually, to the predetermined loading level, followed by the start of the thermal cycle under load on cooling. During the loading (and later, during unloading), the strain gage and thermocouple outputs were registered continuously. If the run was an unloaded one (with the lower grips retaining in position, contributing a negligible external stress of 10.4 psi), then the thermal cycle on cooling was initiated immediately after reaching 85°C.

For all runs, the cooling was intended to be continued down to the regime of temperatures where the circumferential strain gage output showed a linear decrease as a function of decreasing temperature, corresponding to the point where the major part of the sample was transformed to the low temperature martensite phase. When the low limit temperature was achieved (or the experimentally available lowest temperature), the cycle direction was reversed, followed by sample heating again to the high temperature limit. At this point, the sample was unloaded gradually, the loading tray was detached, and the run was completed.

The first run was carried out under zero applied load, followed by a sequence of runs, where in every second run the applied stress level was increased, i.e., for each loaded run, two adjacent unloaded runs were conducted, one before and one after it. The last run was done at an intermediate applied stress between the maximum (10.8 ksi) and the minimum (zero load) applied stresses, in order to investigate the effect of the cycling history. For sample P8-1, the loading sequence of thermal cycling is given in Table 2.

Table 2. Chronological events of cycles performed on sample P8-1, indicating extreme temperature limits during cycling, approximate applied uniaxial stress levels, and temperature fluctuations during loading and unloading

Run	Starting Temp [°C]	T_{max}[a] [°C]	T_{min} [°C]	T_{max}[b] [°C]	Unloading ΔT [°C]	Finishing Temp. [°C]	Applied Stress, σ_a[ksi]	Notes
2	22	95	-33	75	-	21	0	first complete cycling
3	23	84	-10	74	+1	26	1.5	
4	23	85	-22	85	-	25	0	
5	23	85	- 6	85	-1	31	3.0	
6	23	85	-22	85	-	85	0	
7	85	83	- 6	85	-1	84	4.5	contd. from run 6
8	84	84	-19	85	-	17	0	contd. from run 7
9	23	85	-30	85	+1	86	6.0	
10	86	86	-23	85	-	85	0	contd. from run 9
11	85	86	-18	86	-1	30	7.5	contd. from run 10
12	22	85	- 6	85	-	85	0	
13	85	86	-11	85	0	85	9.0	contd. from run 12
14	85	85	-27	85	-	28	0	contd. from run 13
15	24	85	-27	85	0	85	10.5	
16	85	85	-27	85	-	85	0	contd. from run 15
17	85	84	-31	86	0	35	4.5	contd. from run 16

[a]First upper limit.

[b]Second upper limit.

Upon completion of the sequence given in Table 2, sample P8-1 was stripped of its sensors and annealed at 600°C for 1 hour under vacuum, followed by furnace cooling. The same anneal was given to samples P0-4 and P7-1. Then, the samples were electropolished and new sensors were attached to their surfaces.

Samples P0-4 and P7-1 were cycled in temperature under uniaxial applied tension between 0 and 9.0 ksi, and 0 and 6.9 ksi, respectively. The cycling was carried out in a continuous manner, one cycle after the other, with ever increasing applied stress. Loading took place at 125±1 and 145±1°C for sample P0-4 and P7-1, respectively. Also, the minimum temperatures during the cycling for P0-4 and P7-1 were kept constant at -25 and -30°C, respectively.

Sample P8-1 (now in the annealed condition) was cycled first under constant uniaxial tensile stress for a given cycling run. The stresses were increased progressively from run to run, starting at zero stress and incerasing to 7.8 ksi. Then, following the same pattern, still under tension, the applied stress was decreased back to near zero. At this stage, the sample was transferred to the compression device, and compressed through constant compression cycles down to -7.6 ksi. These runs were followed by runs where the compression load was reduced in the reverse sequence back up to near zero compression. For all runs, the limiting temperatures were constant at 90 and -40°C. All loadings and unloadings were done at the high temperature limit.

By the above described experimental procedure, strain-temperature loops were generated for the various constant values of applied stress. All the results were registered as digital printouts. These results were further analyzed using a Commodore Pet 2001 digital microprocessor.

3. RESULTS AND DISCUSSION

3.1 Thermal Cycling of Sample P8-1 in the As-Fabricated Condition Under Applied Tensile Stress

3.1.1 Presentation of results

The results of thermal cycling experiments on sample P8-1 are presented in graphical form. The strain-temperature loops for axial and circumferential strains were generated by the data acquisition system directly from the experimental data points as they were registered by the data logging system. In Fig. 20 and following figures, each point (+) corresponds to a measured value. These points were connected by a solid line drawn in later manually.

For each strain-temperature plot, the abscissa is the temperature axis in °C. As the format of the plotter is constant, this means that the scale may change from run to run, but it remains the same for a given run for both axial and circumferential strains. The ordinates, which are the strain outputs, vary in units according to the maximum strain observed for the run and strain gage of interest, and are given either in units of % strain or microstrain.

Each loop is labelled by arrow heads to designate the direction of the thermal cycle (whether cooling or heating). The starting point of the cycle (S) and the finishing point (F) are also indicated. When the run was under loaded condition, the strain corresponding to loading and unloading was indicated by (L) and (U), respectively.

Figs. 20(a) and (b) show the axial and circumferential strains, respectively, upon thermal cycling of sample P8-1 under zero applied load. Figs. 21 through 26 show the axial and circumferential strains under various applied uniaxial loads.

3.1.2 The use of differential curves in the characterization of strain-temperature hysteresis loops

Differential strain-temperature curves were generated for two reasons to determine the characteristic temperatures, such as the start and finish temperatures M_S, M_F, A_S and A_F, and to investigate the strain hysteresis loop shape upon phase transformation. These curves are shown for the differential axial and circumferential strains of sample P8-1 in Figs. 27-33, corresponding to the runs shown in Figs. 20-26. The differential curves were generated as described in Appendix A.

In the first complete zero-load run (run 2) shown in Fig. 27, the negative peaks of the differential strain curves were found to be relatively smooth for both the axial and the circumferential directions, indicating singly activated processes. The differential peaks for the axial and circumferential strains are in the same general direction, and for both cases the peak generated on heating is sharper and has a higher peak value than those generated on cooling. It appears qualitatively that the total areas under the differential peaks on cooling and heating are equal.

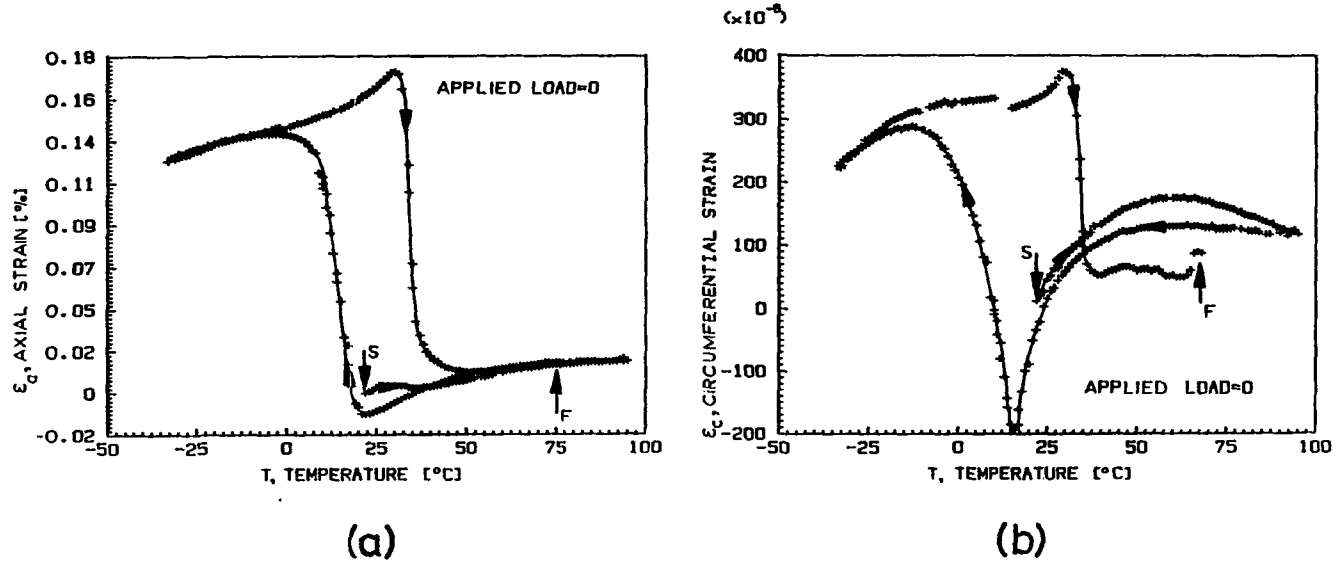


Figure 20. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under zero applied stress during run 2

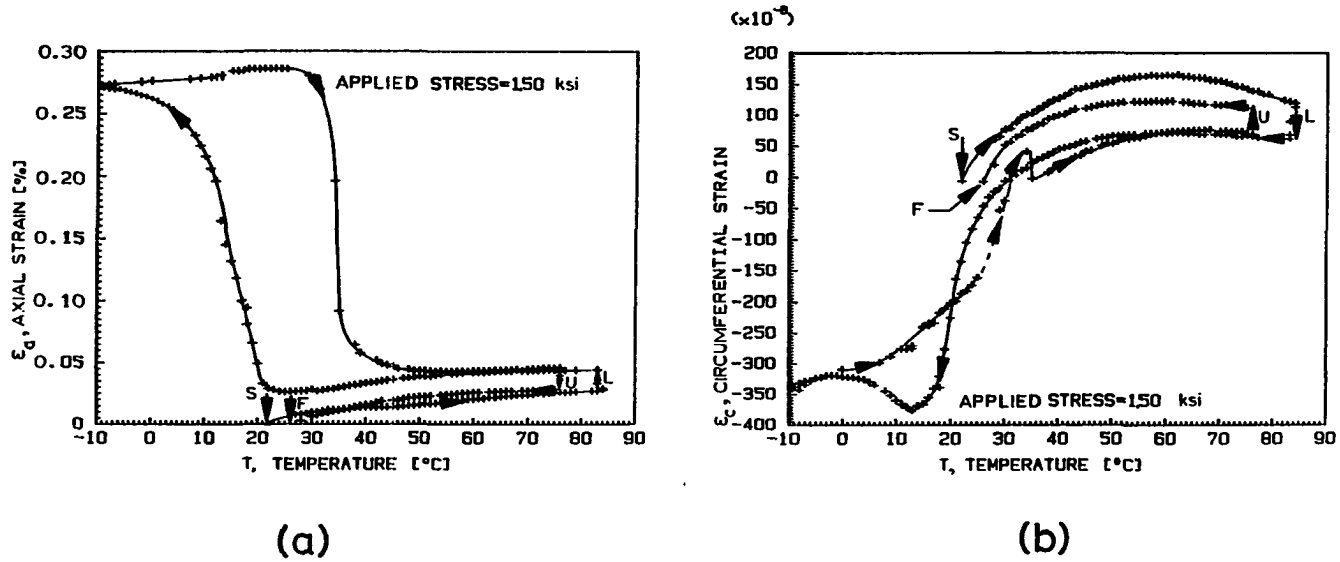


Figure 21. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under 1.50 ksi applied stress

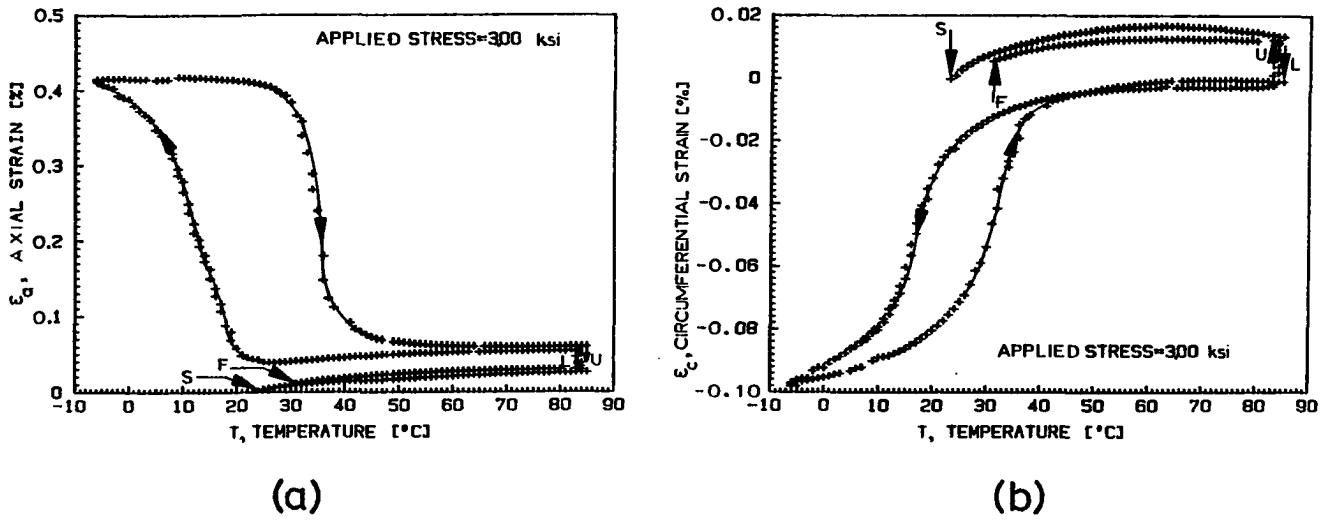


Figure 22. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under 3.00 ksi applied stress

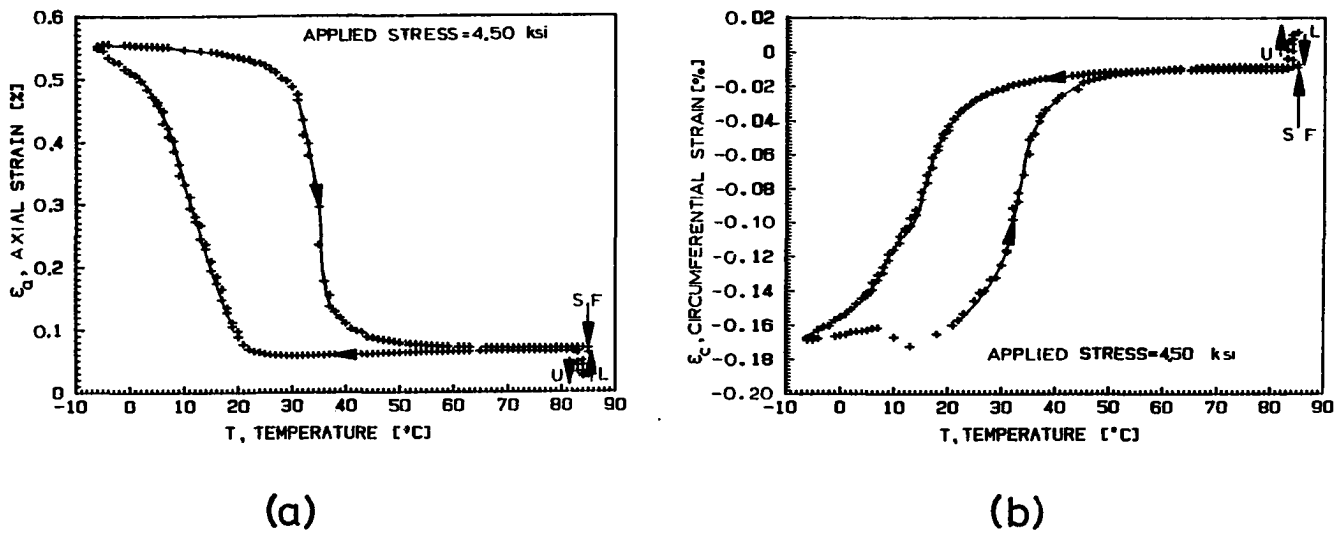


Figure 23. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under 4.50 ksi applied stress

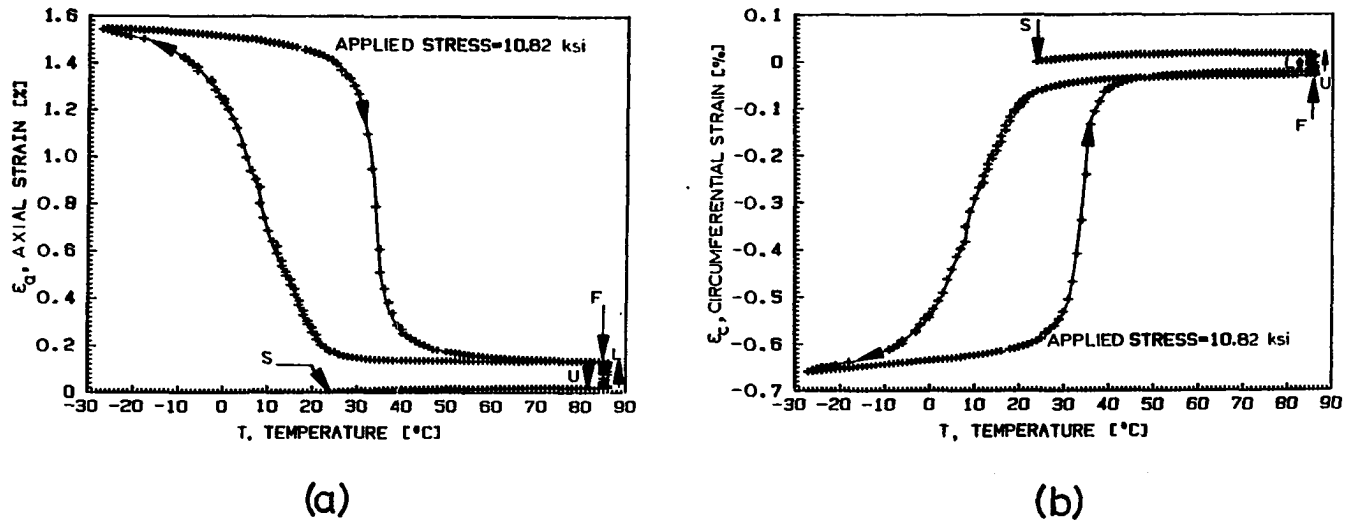


Figure 24. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under 10.82 ksi applied stress

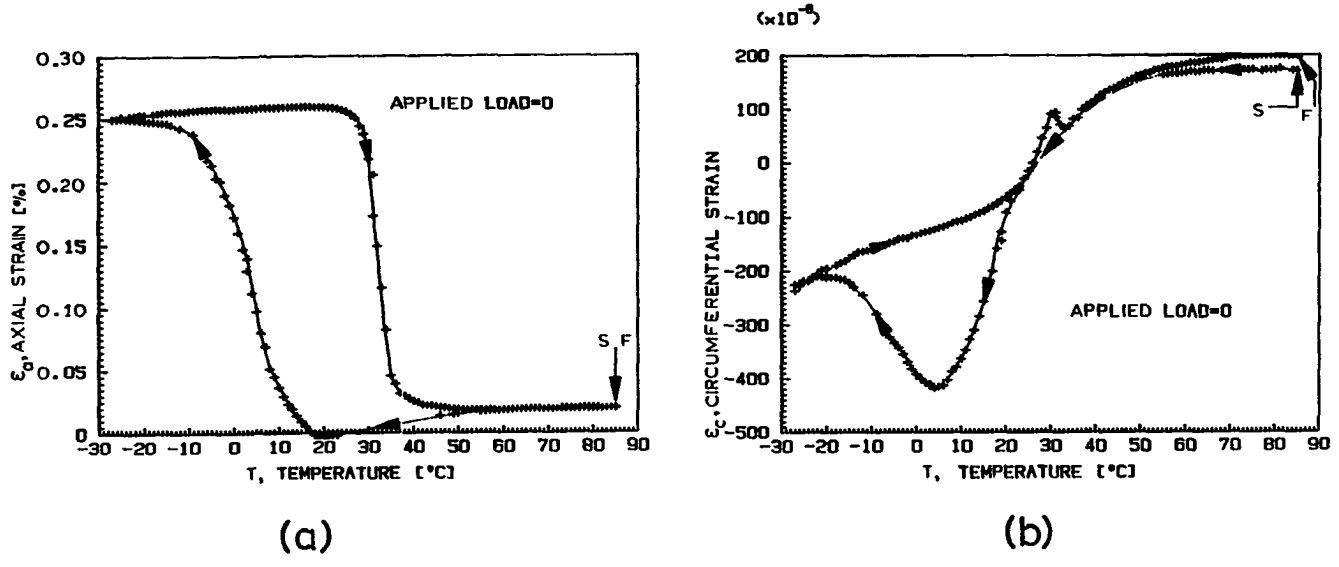


Figure 25. Strain hysteresis loops for (a) axial and (b) circumferential strains as a function of temperature of sample P8-1 under approximately zero (actually, 10.4 psi) applied stress for run 16

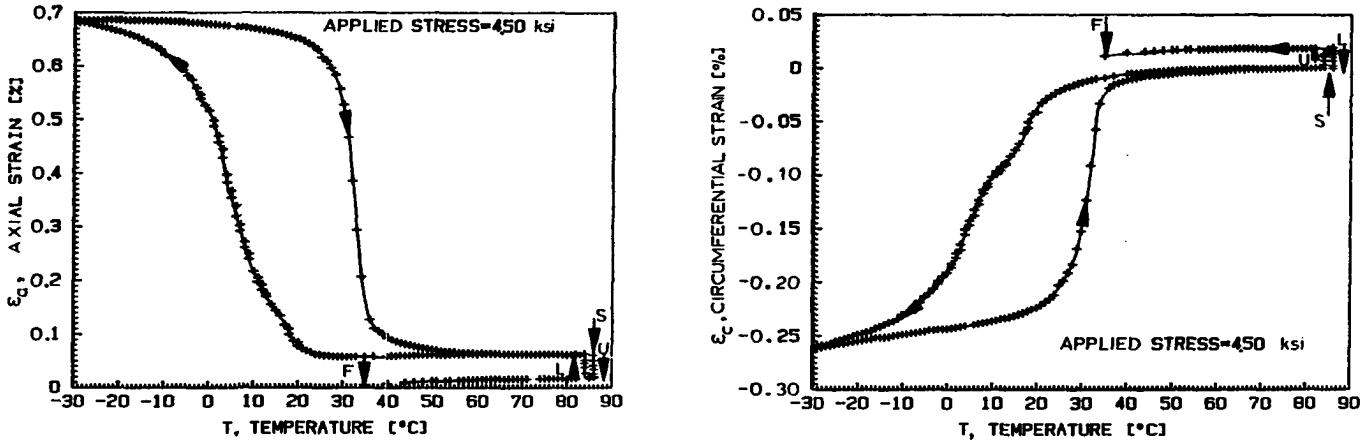


Figure 26. Strain hysteresis loops for (a) axial and (b) circumferential strain as a function of temperature of sample P8-1 under 4.50 ksi applied stress for run 17

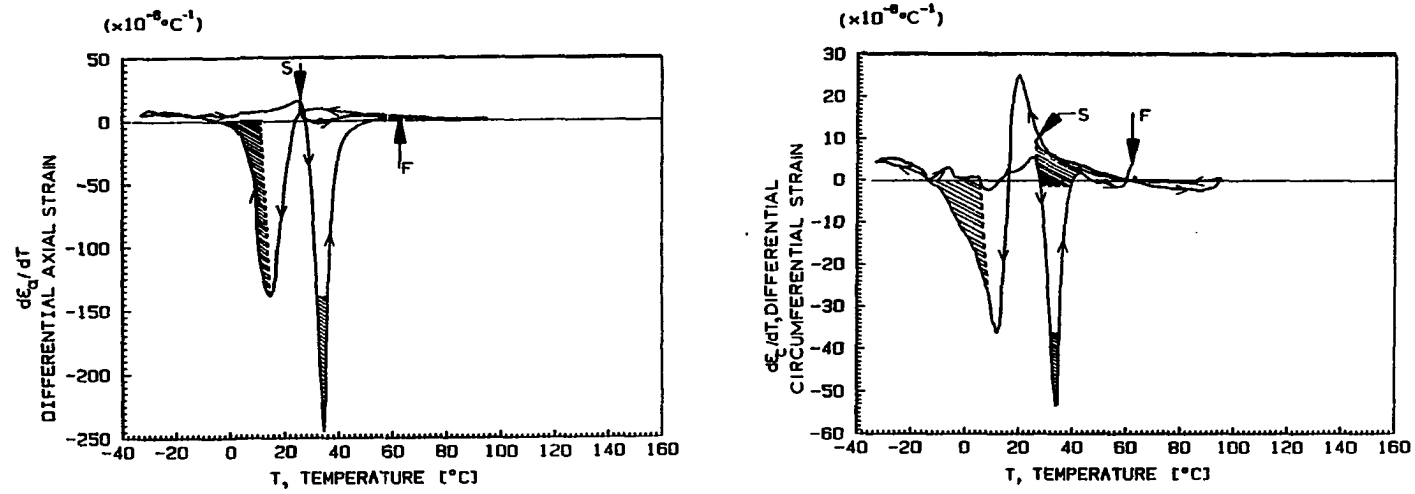


Figure 27. Differential of the (a) axial (Fig. 20(a)), and (b) circumferential (Fig. 20(b)) strain hysteresis loops as a function of temperature of sample P8-1 under zero load

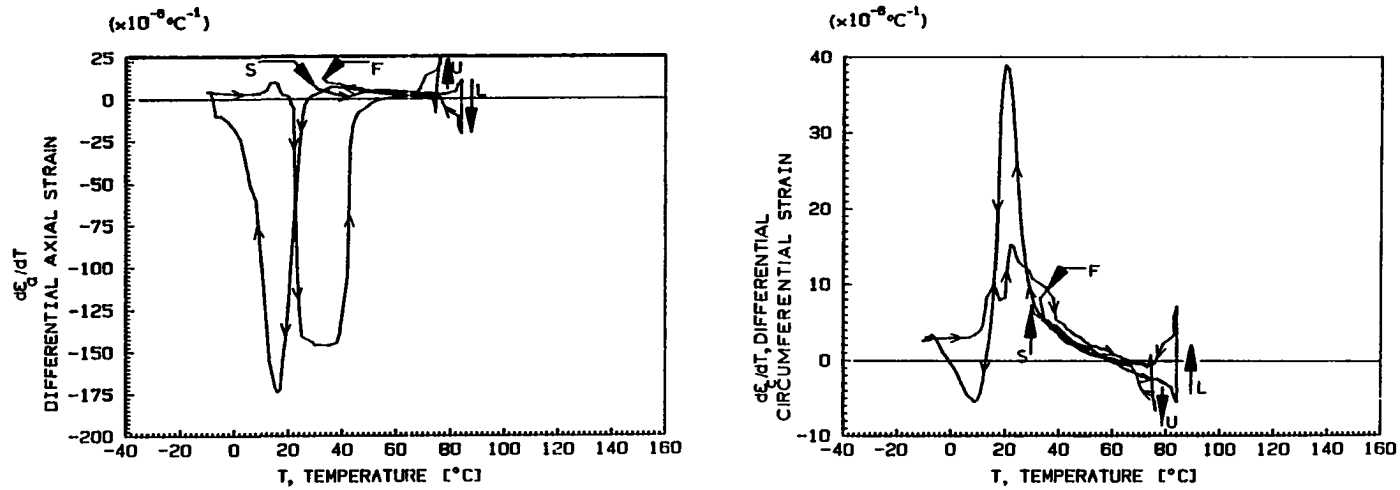


Figure 28. Differential of the (a) axial (Fig. 21(a)), and (b) circumferential (Fig. 21(b)) strain hysteresis loops as a function of temperature of sample P8-1 under 1.50 ksi applied stress

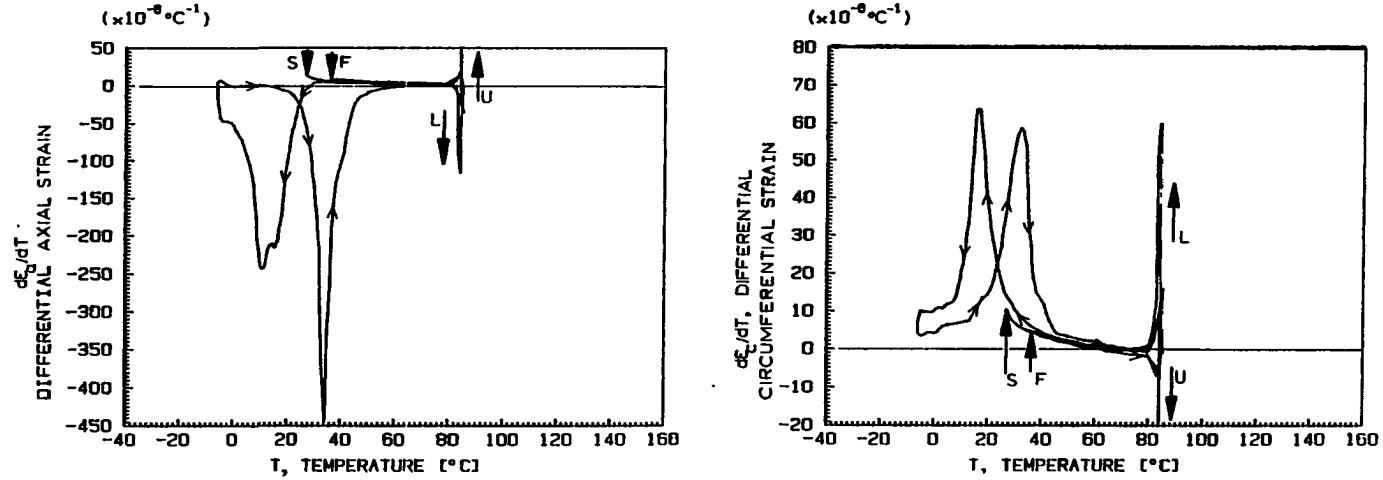


Figure 29. Differential of the (a) axial (Fig. 22(a)), and (b) circumferential (Fig. 22(b)) strain hysteresis loops as a function ot temperature of sample P8-1 under 3.00 ksi applied stress

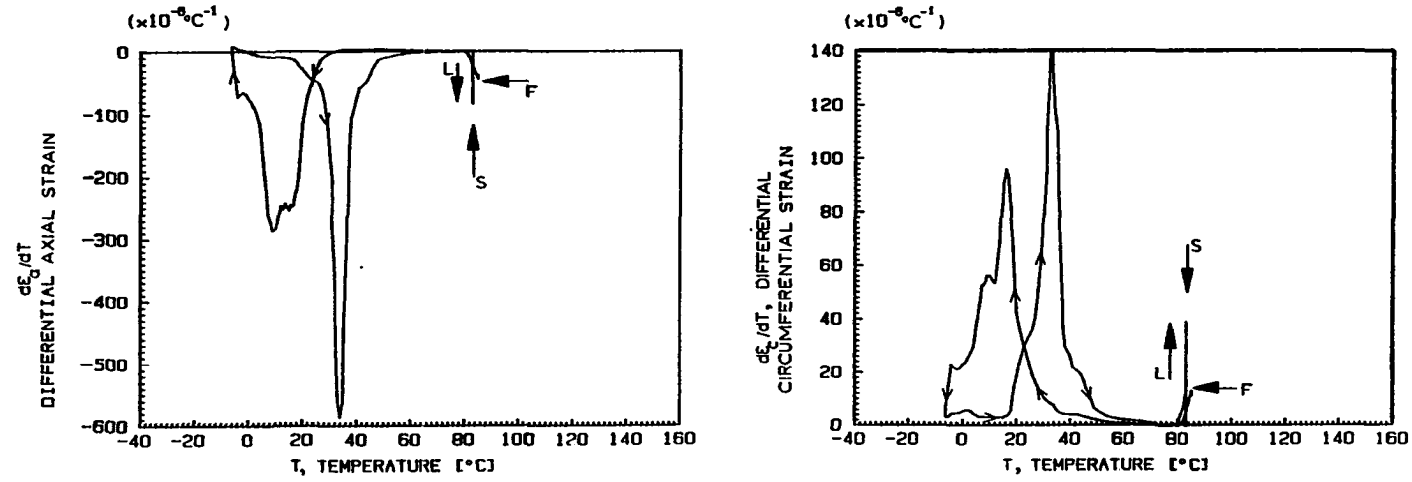


Figure 30. Differential of the (a) axial (Fig. 23(a)), and (b) circumferential (Fig. 23(b)) strain hysteresis loops as a function of temperature of sample P8-1 under 4.50 ksi applied stress

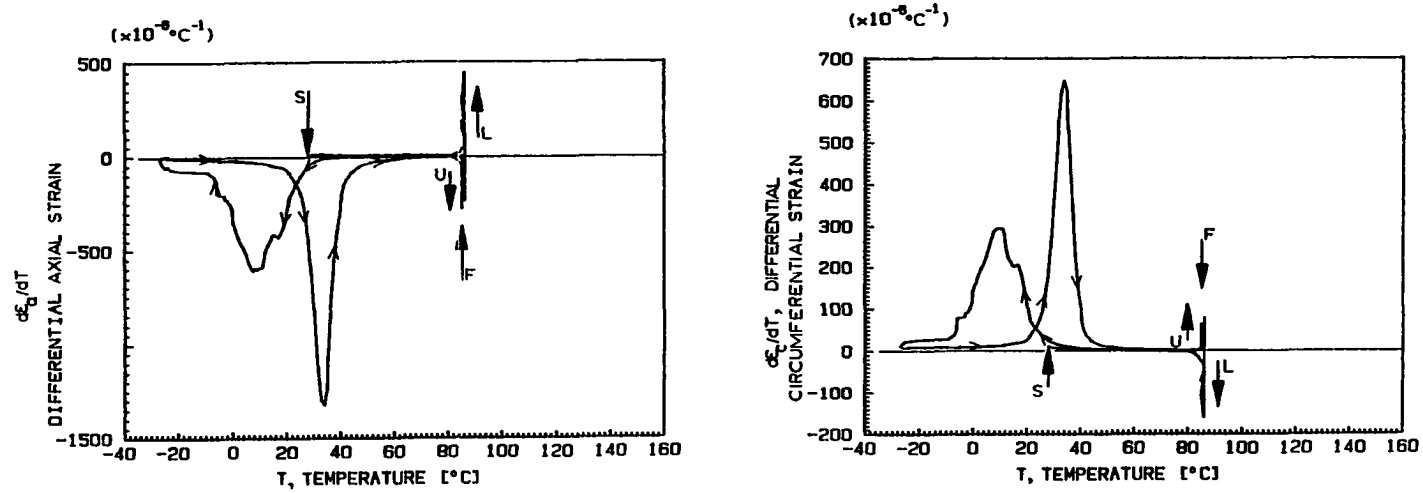


Figure 31. Differential of the (a) axial (Fig. 24(a)), and (b) circumferential (Fig. 24(b)) strain hysteresis loops as a function of temperature of sample P8-1 under 10.82 ksi applied stress

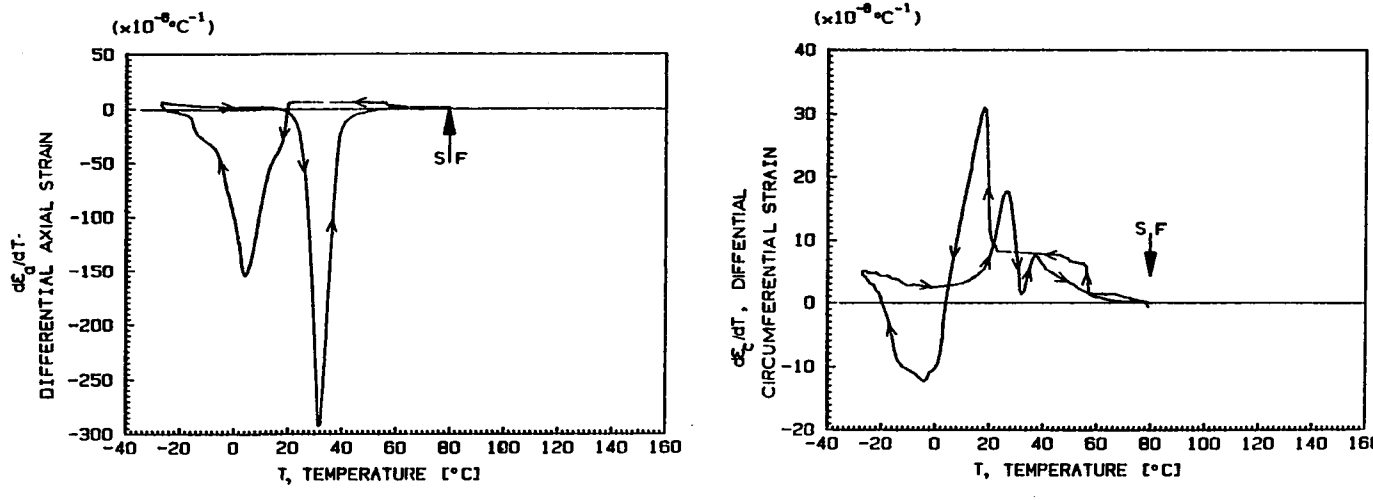


Figure 32. Differential of the (a) axial (Fig. 25 (a)), and (b) circumferential (Fig. 25(b)) strain hysteresis loops as a function of temperature of sample P8-1 under approximately zero (actually 10.4 psi) applied stress for run 16

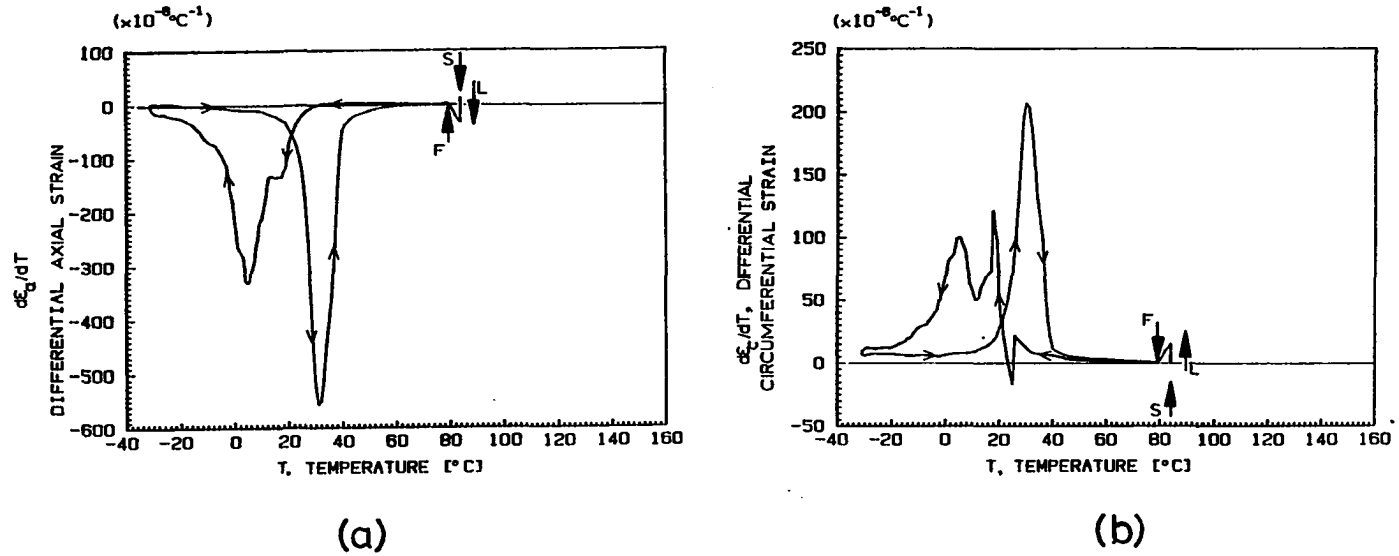


Figure 33. Differential of the (a) axial (Fig. 26(a)), and (b) circumferential (Fig. 26(b)) hysteresis loops as a function of temperature of sample P8-1 under 4.50 ksi applied stress, following the highest applied stress cycle (run 15) and its adjacent nearly zero load run (run 16)

For zero-load differential strain curves following runs of increasing applied stresses (runs 4, 10, 14, and 16)[1] there are several progressively developing features:

(1) For the differential axial strain:
   (a) On the cooling peak, a secondary peak develops with increasing run number.
   (b) The heating peak becomes more and more symmetric around its maximum value with increasing run number.

(2) For the differential circumferential strain:
   (a) The positive peak increases in height and area, both on cooling and on heating, with increasing run number.
   (b) The corresponding negative peak decreases in height and area with increasing run number and more so on heating than on cooling.
   (c) For run 10 (following a 6.0 ksi applied stress run), the peak values for the negative peaks on cooling and on heating become equal. For run 16 (following the 10.82 ksi applied stress run), the negative peak on heating diminishes compared to the other peaks present (Fig. 32(b)).
   (d) For run 10, the positive peak on cooling begins to develop a secondary peak, similar to that for the axial strain peak on cooling. This culminates in Run 16 in a wide plateau from 60°C to 20°C (Fig. 32 (b)).

---

[1]The results for run 16 only is shown in a figure (Fig. 32).

The runs followed a sequential pattern, where each loaded run was followed by a zero-stress one, with ever increasing stresses from one loaded run to the next one (see Table 2). The above features of the differential strain curves, namely 1(a), 1(b), and 2(a) through 2(d), should be attributed to the thermomechanical history of the sample during cycling. This and related effects will be called "training" from here on.

Training can be described as follows. For the circumferential strain, on cooling the sample from the HTP under no load (run 16) and neglecting the plateau, the sample begins to behave as if under load, i.e., it shows a high positive peak, as it did under the previous, high applied stress run (compare Figs. 32(b) and 31(b)). When the MPT proceeds near T_C, as defined in Appendix A, the sample begins to behave again as it did at zero load before training, i.e., it begins to exhibit the negative peak characteristic of the zero-load run (compare Figs. 32(b) and 27(b)). In addition, on heating from the LTP, the sample behaves at first as it did on the first run under no load prior to training. But, instead of developing a big negative peak as at 35°C in Fig. 27(b), it shows only a small drop at 35°C, as seen in Fig. 32(b). The cycle after training is terminated by a weak positive peak near 40°C (Fig. 32(b)) at the location where it was for the previous run under load (Fig. 31(b)). Little or no evidence of this positive peak was present before training (Fig. 27(b)).

The training effect is especially strong for the axial strain. With the exception of the plateau (Fig. 32(a)), the sample behaves as it did under load (Fig. 31(a)), although with smaller peak values. Comparing the zero-load run after training (Fig. 32(a)) with the zero-load run before training (Fig. 27(a)), we see the same pattern of peaks (negative peaks at about 10°C on cooling and 30°C on heating), but the peak values are increased.

The differential curves for the loaded runs from 3 ksi applied stress up to above 10 ksi are very different from those of the zero load runs. Each curve contains a major peak on cooling and a major peak on heating (see, for example, Fig. 31). The peaks for the axial strain are negative, whereas for the circumferential strain they are positive, i.e., no longer in the same direction for axial and circumferential strains as was the case for zero-load runs, as shown, for example, in Fig. 27.

The peak on heating always has a greater magnitude than on cooling. The heating peak is smooth and more symmetric than the cooling peak (e.g., Fig. 31), indicating that the phase transformation on heating is singly activated. On the other hand, the axial cooling peak begins to develop a secondary peak at and following the 3 ksi applied stress run (e.g., Fig. 29(a)). The circumferential cooling peak develops a secondary peak at and following the 4.5 ksi applied stress run (e.g., Fig. 30(b)). The main and secondary peaks become more and more separated as the stress increases, as is seen by comparing Figs. 30(b) and 31 (b).

Altogether, the transformation on cooling under uniaxial applied stress does not appear to be singly activated.

Run 3 at 1.50 ksi stress (Figs. 28(a) and (b)) is a special case among the loaded cycles. To begin with, the axial differential strain on heating has a truncated peak. The differential circumferential strain behaves on cooling and heating somewhat similarly to a zero-load run after training, i.e., it shows on cooling a large positive peak, followed by a small negative peak and on heating a generally positive fluctuating peak (compare Figs 28(b) and Fig. 32(b)). This parallelism is even more striking when the integral circumfential strain hysteresis loops are compared (Figs. 21(b) and 25(b)), in that both of them are twisted. Since the strain hysteresis loops are characteristic of the MPT, it is likely that one mechanism underlies both behaviors.

For run 17 at 4.5 ksi (Fig. 33), the differential axial strain is that of a normal loaded run (Fig. 33(a)). On the other hand, the differential circumferential strain on cooling has an extra negative transient at 25°C, followed by a double positive peak indicating that an instability exists in this temperature range.

It is noted here that the common factor to the cases shown in Figs. 21 (b) and 25(b) is training, although viewed from two extreme limits. To be more specific, it was found that in wires that were trained by a large number of transformation fatigue cycles (of the orders of milions), the differential axial strain showed a splitting on cooling, but not on heating [86]. Such a splitting, or at its early stages the secondary

peak development, occurs in sample P8-1, even after a few cycles, provided they are carried out under applied load. Hence, the training seems to be connected to the applied stress and not to the high number of fatigue cycles.

The differential strain behavior with increasing applied load is summarized in Fig. 34. Each differential strain ordinate is scaled to the same units, which enables an easy comparison of the changes in the peak values on cooling and on heating with applied stress. It is evident in Fig. 34 that the peak magnitudes increase monotonically with increasing stress. The development of secondary cooling peaks is also clearly observable. When the axial differential strain curves are laid one upon the other as in Fig. 35, their similarity in general shape and the ever increasing area under the peaks with increasing applied load are evident.

Following the procedure outlined in Appendix A, the characteristic temperatures and their corresponding strains were evaluated. Numerical values are given in Table 3. In order to determine the influence of the applied stress on the characteristic temperatures, some of these temperatures are indicated in Fig. 36, together with the shift from their position for zero applied stress. The secondary peak temperatures are designated as T_R (suggesting the possibility that the origin of the peaks is in the B2→R transformation [87]). All the other notations are defined in Appendix A.

Figure 36 and Table 3 show the following overall effects of applied tensile stress on the characteristic temperatures: (1) A_F and M_S are

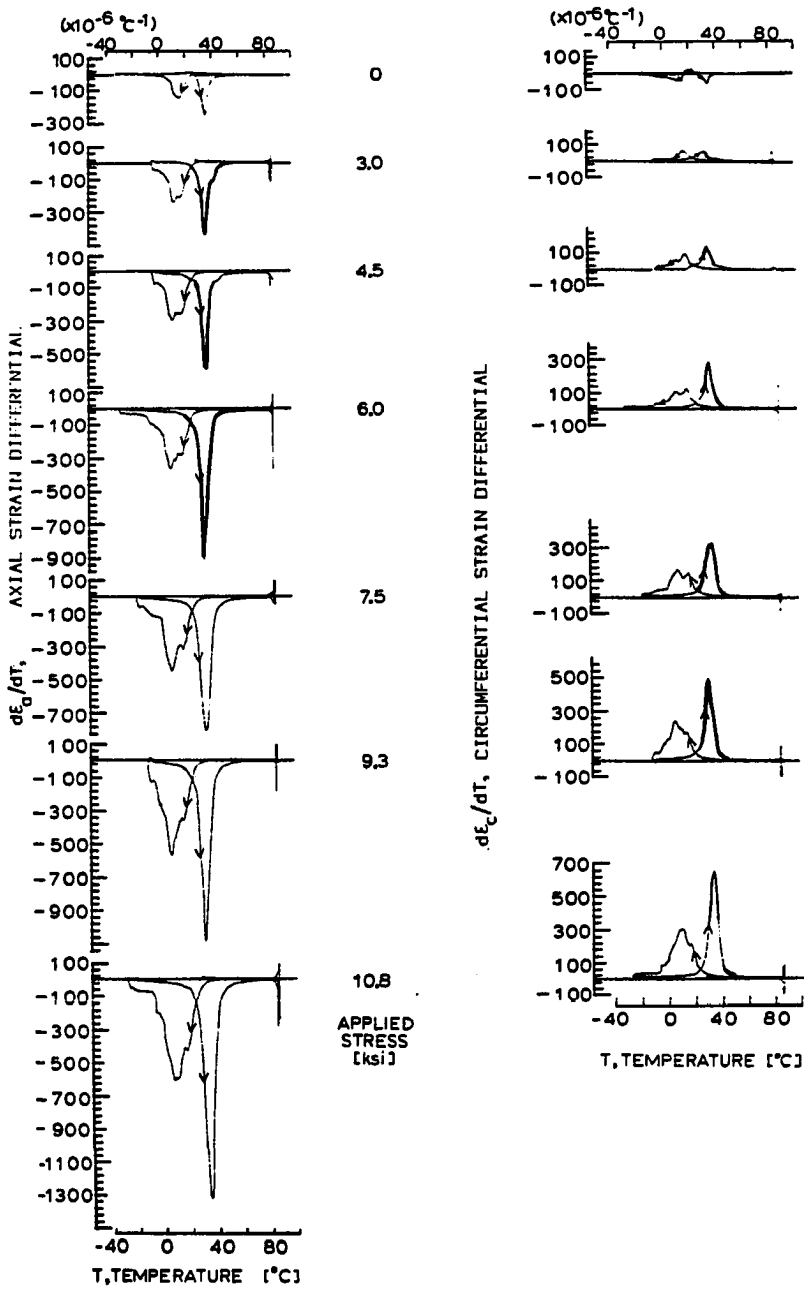


Figure 34. Summarizing differential strain curves as a function of applied stress, from 0 (uppermost curve) to 10.82 ksi (bottom curve), showing increase in peak values on both heating and cooling, and development of secondary peak on the cooling part of the thermal cycle.

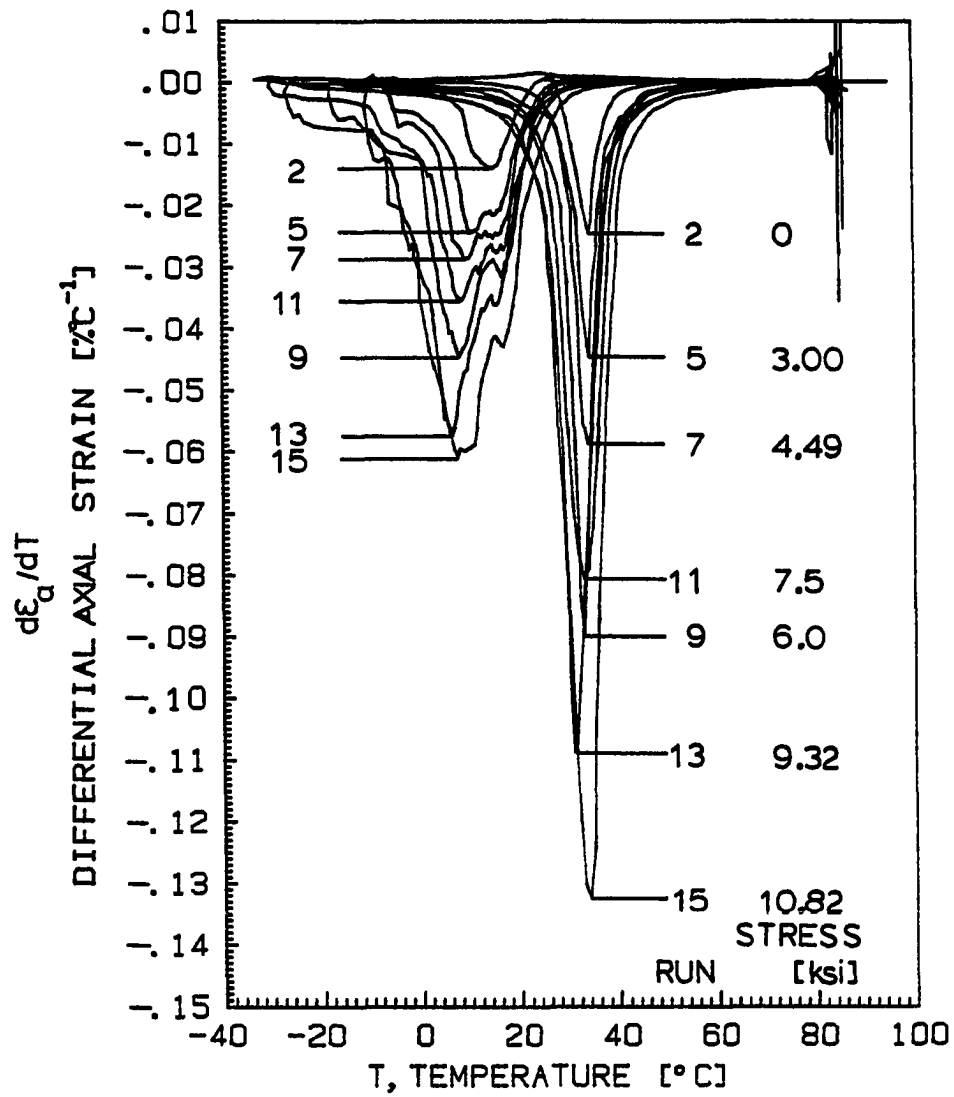


Figure 35. Summarizing curve of the differential axial strains on heating (the right side family of curves) and on cooling (the left side family of curves). Increase in peak values in general, and the indication of multiple activated process development on cooling are clearly visible. Each peak is labelled by means of a horizontal line at its peak value to which its run number is attached. The rightmost column of numbers is the corresponding applied stress in ksi

Table 3. Summarizing data for sample P8-1. All temperatures in [°C], derived from the differential curves. All strain values are in [%], derived from integral curves using the previously circumferential strains. (*) means, σ_a=10.4 psi. (†) means deduced from Figs. 54 and 55, below

σ_a[ksi]	A_S	$\varepsilon(A_S)$	A_F	$\varepsilon(A_F)$	M_S	$\varepsilon(M_F)$	M_F	$\varepsilon(M_F)$	T_H	$\varepsilon(T_H)$	T_C	$\varepsilon(T_C)$	T_E	$\Delta\varepsilon_{T_E}$
0	30	.169	43	.013	21	-.011	7	.129	35	.060	16	.024	26	.154
	28	.036	--	---	24	-.001	-6	.027	35	.012	9	.003	28	.031
1.50	26	.285	46	.045	20	.049	-6†	---	33	.250	12	.195	26	.260
	10	-.028	45	---	33	.004	0	-.036	27	-.013	20	-.022	26	-.010
0*	25	.226	47	.024	17	.030	-2	.198	36	.083	11	.120	25	.025
	29	.034	--	---	28	.002	-5	.021	34	.009	9	-.012	--	---
3.00	28	.398	51	.066	25	.039	<-6†	---	36	.175	12	.234	26	.369
	9	-.090	45	-.006	27	-.017	<-6†	---	33	-.032	16	-.060	23	-.052
0*	26	.241	46	.023	20	-.003	-6	.220	35	.110	9	.129	22	.244
	--	---	--	---	27	.000	-8	.021	--	---	0	-.020	--	---
4.50	23	.530	44	.090	23	.068	<-6†	---	36	.179	10	.330	4	.459
	--	---	--	---	31	-.021	<-6†	---	32	-.102	15	-.093	--	---
0*	26	.253	40	.033	21	.000	-6	.231	35	.091	8	.138	21	.253
	--	---	--	---	30	.002	-13	.018	--	---	7	-.019	--	---
6.00	26	.708	44	.103	25	.068	-8	.722	34	.305	9	.432	24	.659
	32	-.243	43	-.023	25	-.035	-11	-.258	34	-.098	16	-.090	24	-.200

Table 3 (Continued)

σ_a[ksi]	A_S	$\varepsilon(A_S)$	A_F	$\varepsilon(A_F)$	M_S	$\varepsilon(M_F)$	M_F	$\varepsilon(M_F)$	T_H	$\varepsilon(T_H)$	T_C	$\varepsilon(T_C)$	T_E	$\Delta\varepsilon_{T_E}$
0*	28	.241	44	.015	23	-.010	-8	.225	34	.147	6	.133	21	.255
	23	.021	44	.008	30	.004	-12	.009	34	.010	4	-.020	28	.023
7.50	25	.880	49	.110	24	.092	-12	.932	35	.340	8	.550	24	.800
	21	-.335	44	-.030	42	-.025	-8	-.349	35	-.115	8	.225	24	-.282
0*	26	.225	35	.035	20	-.005	--	---	31	.130	6	.107	20	.241
	--	---	34	.007	32	.008	--	---	31	.007	-3	-.015	27	.018
9.32	22	1.080	43	.165	22	.145	<-11†	---	32	.608	3	.900	22	.935
	22	-.435	42	-.037	22	-.055	<-11†	---	32	-.200	10	-.218	22	-.380
0*	24	.260	42	.021	24	.004	-11	.245	32	.075	2	.149	20	.265
	17	.009	43	.013	30	.008	-14	-.005	31	.008	-11	-.011	24	.010
10.82	26	1.365	47	.183	26	.150	-18	1.500	35	.510	8	.820	24	1.245
	23	-.600	43	-.050	26	-.060	-8	-.610	35	-.213	8	-.380	24	-.670
0*	28	.245	42	.020	22	-.005	-10	.240	32	.135	5	.098	23	.060
	22	-.005	53	.017	--	---	-16	-.023	32	.007	-8	-.030	--	---
4.50	19	.652	40	.090	24	.060	-14	.645	33	.341	4	.370	21	.575
	19	-.228	40	-.014	36	-.010	-17	-.246	32	-.116	4	-.155	22	.190

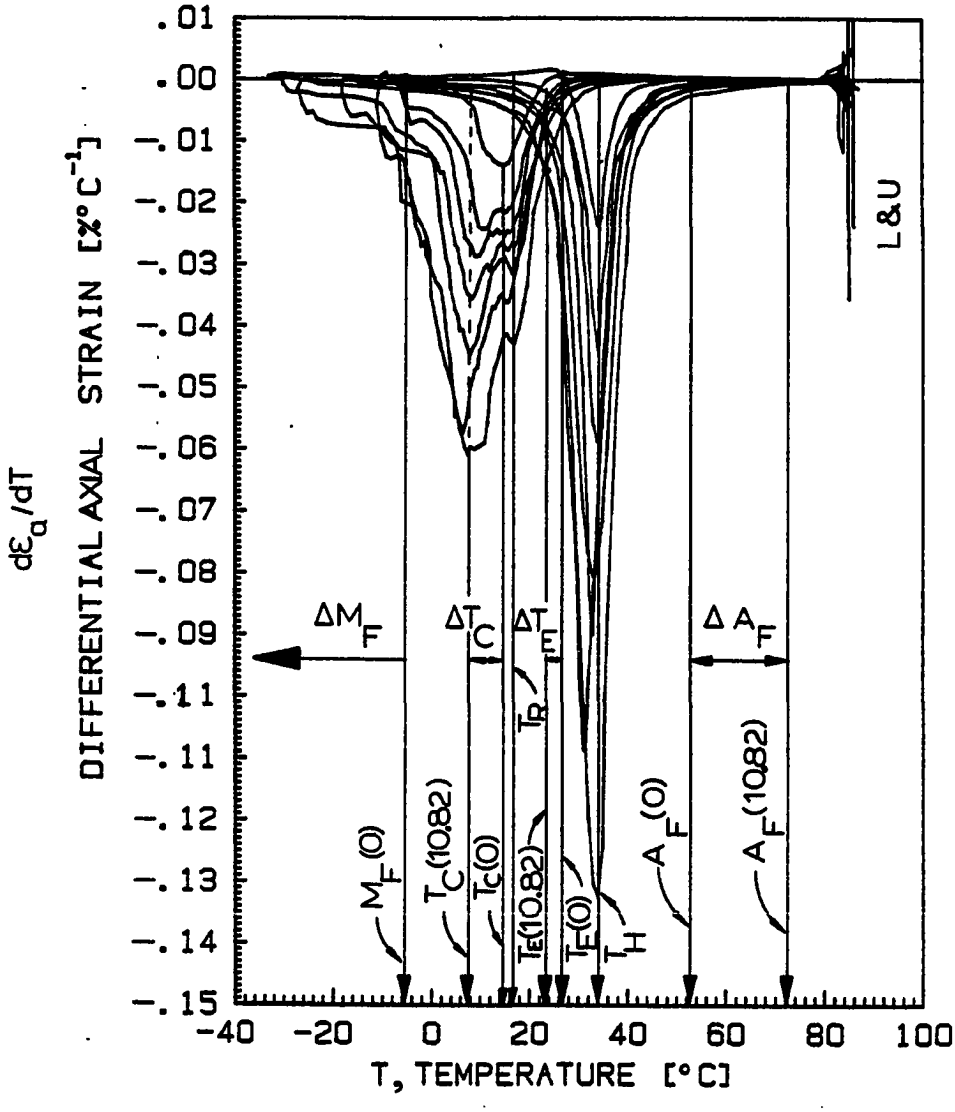


Figure 36. The same summarizing figure as Fig. 35, but showing characteristic temperatures and their shifts due to changing the applied stress. Numbers in parentheses are stress levels in ksi units. ΔM_F, ΔT_C and ΔT_E are negative, ΔT_H and ΔT_R are zero, ΔM_S (not shown) and ΔA_F are positive. The most pronounced shift occurs for M_F

increased; (2) T_C T_E, and M_F are decreased; and (3) T_H and T_R are unaffected. Let us consider these changes in characteristic temperatures in terms of the Gibbs free energy balance underlying the transformation. The differential of the free energy attendant upon the formation of a differential volume, dV of martensite and interfacial area dA between the martensite and austenite may be given by

$$dG = -g_V dV + g_A dA + \left(\sum_{p,q} \int \sigma_{pq}^{(i)} d\varepsilon_{pq} \right) \cdot dV - \sigma^{(a)} \Delta\varepsilon_a dV \quad (4)$$

Here, g_V is change in free energy per unit volume, which serves as the chemical driving force and g_A is the free energy change per unit surface area at the austenite-martensite interface. The interfacial energy associated with twin bands within the martensite plate is neglected in equation (4). The classical representation of g_V is given in Fig. A9 of Appendix A. The martensite product is known to form by an invariant plane strain, which is illustrated in Fig. B5 of Appendix B as the distortion that takes ABCD into AB"C"D. This distortion is resisted by volume and shear constraints, which result in internal stresses $\sigma_{pq}^{(i)}$ and strains $\varepsilon_{pq}^{(i)}$. The corresponding strain energy is given by the third right-hand term in equation (4). Finally, the last term in equation (4) represents the decrease in free energy due to the axial transformation strain $\Delta\varepsilon_a$, which tends to relax the constant applied stress $\sigma^{(a)}$. The increases in A_F and M_S with increasing $\sigma^{(a)}$ indicate an increase in the stability of the martensite. This can be understood in terms of the increase in the magnitude of the last term in equation (4) with increasing

$\sigma^{(a)}$ due to two factors: the increase in $\sigma^{(a)}$ itself and the associated increase in $\Delta\varepsilon_a$ as, with increasing $\sigma^{(a)}$, more and more martensite variants form that produce strains that tend to be aligned along the uniaxial stress direction. An increase in M_S with increasing $\sigma^{(a)}$ was also observed by Hsu [80] for Ni-Ti wires. After transformation fatigue cycling into the millions of cycles, the effect of $\sigma^{(a)}$ on M_S was greatly reduced, presumably because of the internal stresses produced upon high-level cycling that tend to mask the effect of the applied stress [80].

An examination of Fig. 36 and Table 3 indicates that the decrease in M_F with increasing applied stress is especially pronounced. As more and more of the favorably oriented martensite variants form (i.e., those variants with extensional strains in directions near the stress axis), the internal stresses $\sigma_{pq}^{(i)}$ and strains ε_{pq} tend to build up until the point is reached when less favorable variants must form, whose stresses and strains tend to cancel those due to the more favorable variants. These less favorable variants do not relieve the applied stress as much and their effect on the last term in equation (4) is smaller. This means that $\sigma^{(a)}$ is less effective in reducing G when the less favorable variants are called into play, which in turn necessitates greater undercooling in order to continue the transformation. Hence, M_F decreases with increasing $\sigma^{(a)}$. The same rationale as given above may apply to the observed decrease in T_C with increasing $\sigma^{(a)}$. As explained

in Appendix A, T_C is the peak temperature on cooling (see, for example, the differential strain peak at 8°C for sample P8-1 upon cooling under 10.82 ksi stress, as shown in Fig. A14 of Appendix A). It is reasonable to suggest that the differential strain reaches its peak value when the less favorable variants start operating. For reasons described above, T_C should therefore decrease with increasing $\sigma^{(a)}$.

As concerns the observed lack of dependence of T_R and T_H on $\sigma^{(a)}$, we recall first that T_R is considered to be the temperature at which the intermediate rhombohedral phase forms on cooling, as discussed, for example, by Ling and Kaplow [87]. These authors have pointed out that T_R is not affected by prior cycling, which implies that the formation of the rhombohedral phase does not depend on internal stresses. In addition, Hsu [80] found that T_R does not change as a function of applied stress, as deduced from the fact that the anomalous resistivity peak on cooling occurred at the same temperature, regardless of the magnitude of $\sigma^{(a)}$ up to 8 ksi.

The single well-defined peak on heating, characterized by T_H, signals the reverse transformation to the austenite or B2 structure. In analogy to the discussion of T_C and M_F, the observation that A_F increases with increasing $\sigma^{(a)}$ suggests that T_H should increase somewhat also. However, this was not observed.

3.1.3 Magnitudes of axial and circumferential strains and the phase angle between them

Materials that undergo MPT and exhibit thermoelastic behavior are likely to show the features of anelastic behavior, as, for example, caused by interface migration, as suggested by Khachin and Solovev [75]. Any anelastic behavior depends upon a phase shift (in time) between two characteristic parameters. In this section, it is shown that the axial and the circumferential strains exhibit a phase shift, whose magnitude depends upon the applied stress.

Evaluation of a direct, continuous functional relation between ε_c and ε_a is possible by considering temperature as a parameter. Then the curves ε_a versus T and ε_c versus T can be plotted ε_c versus ε_a. This approach is taken in Figs. 37 to 43, for sample P8-1. In each figure, the direction along the path during thermal cycling is indicated with arrowheads together with starting (S), finishing (F), loading (L), and unloading (U) points. Along the path, points are designated in alphabetic sequence. To the right of each designated point, the temperature in °C is given in parentheses. Below each designated point, in square brackets, a pair of numbers appears, separated by comma. Each of these numbers represents the relative volume dilation $\Delta V/V_o$ (in units of 10^{-3}), based on the assumptions that the volume change is given by $\Delta V/V_o=\Delta\varepsilon_a+2\Delta\varepsilon_c$; or by $\Delta V/V_o=3(\Delta\varepsilon_a+\Delta\varepsilon_c)/2$, for the first and second number in parentheses, respectively. The first equation assumes that both transverse strains equal ε_c, and the second equation assumes that the

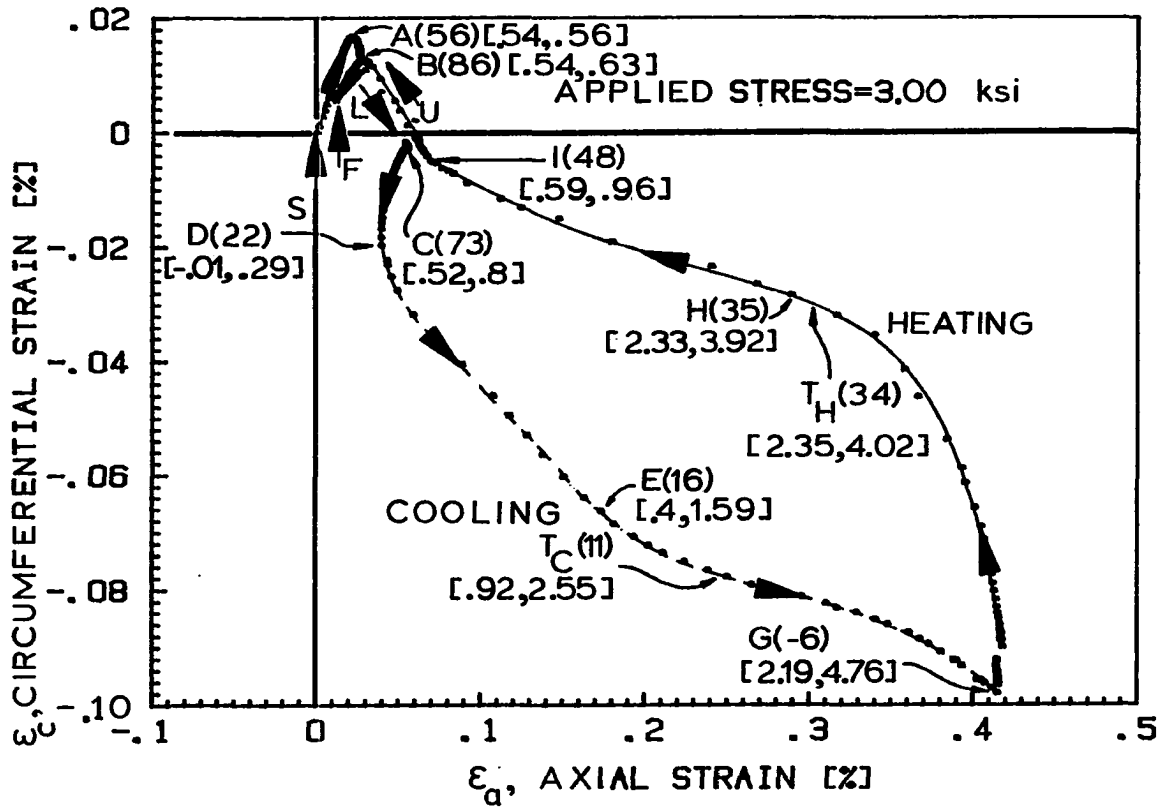


Figure 37. The functional relationship between the circumferential and axial strains for sample P8-1 under 3.00 ksi applied stress. Cooling from C to G (dashed line), and heating from G to near I (full line)

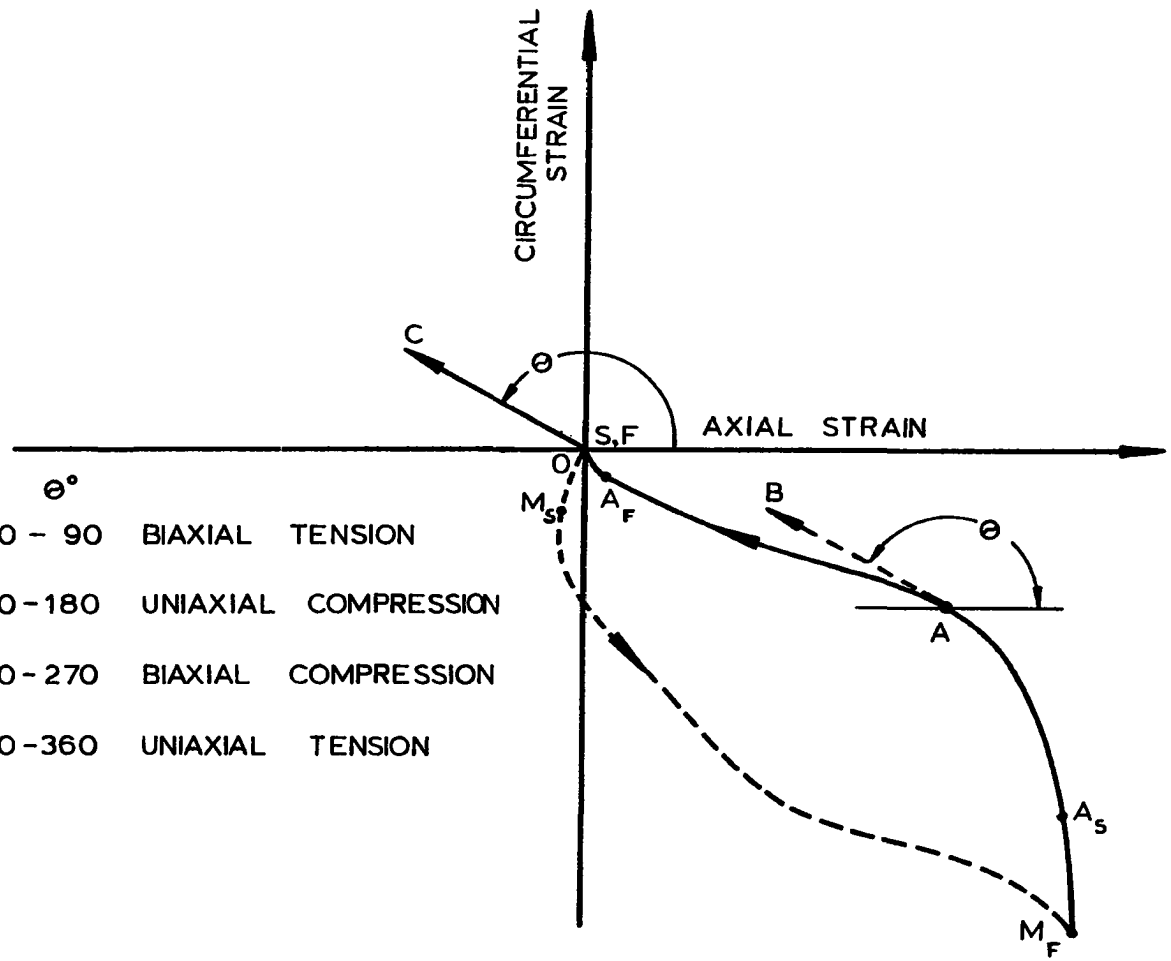


Figure 38. Schematic representation of ε_c-ε_a loop. The strains are zeroed at the start(S) of the run in the HTP. The direction θ of the tangent line (AB) at a given point A is indicative of the total stress state (applied and internal stresses)

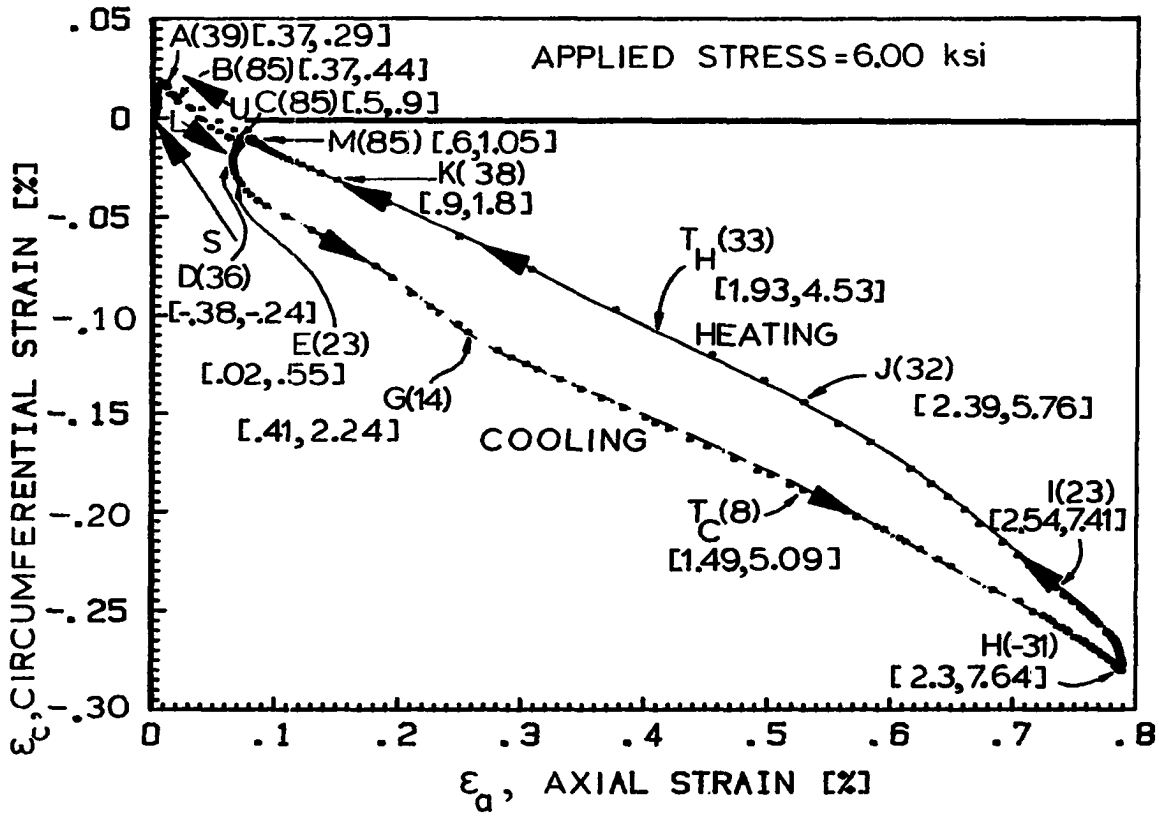


Figure 39. The functional relationship between the circumferential and axial strains for sample P8-1 under 6.00 ksi applied stress. Cooling from C to H (dashed line), and heating from H to M (full line)

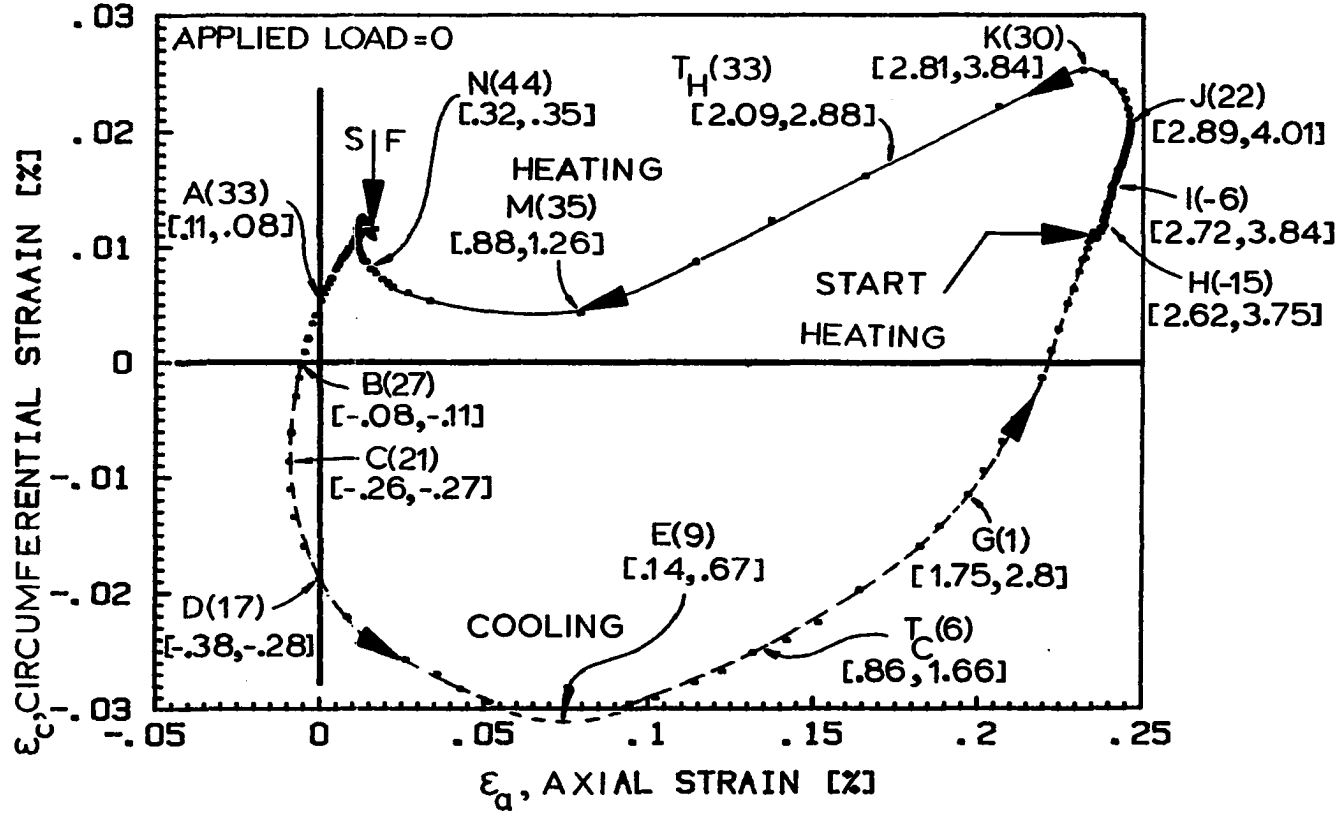


Figure 40. The functional relationship between the circumferential and axial strains for sample P8-1 under no load condition following run 9. Cooling from S to near H (dashed line), and heating from near H to F (full line)

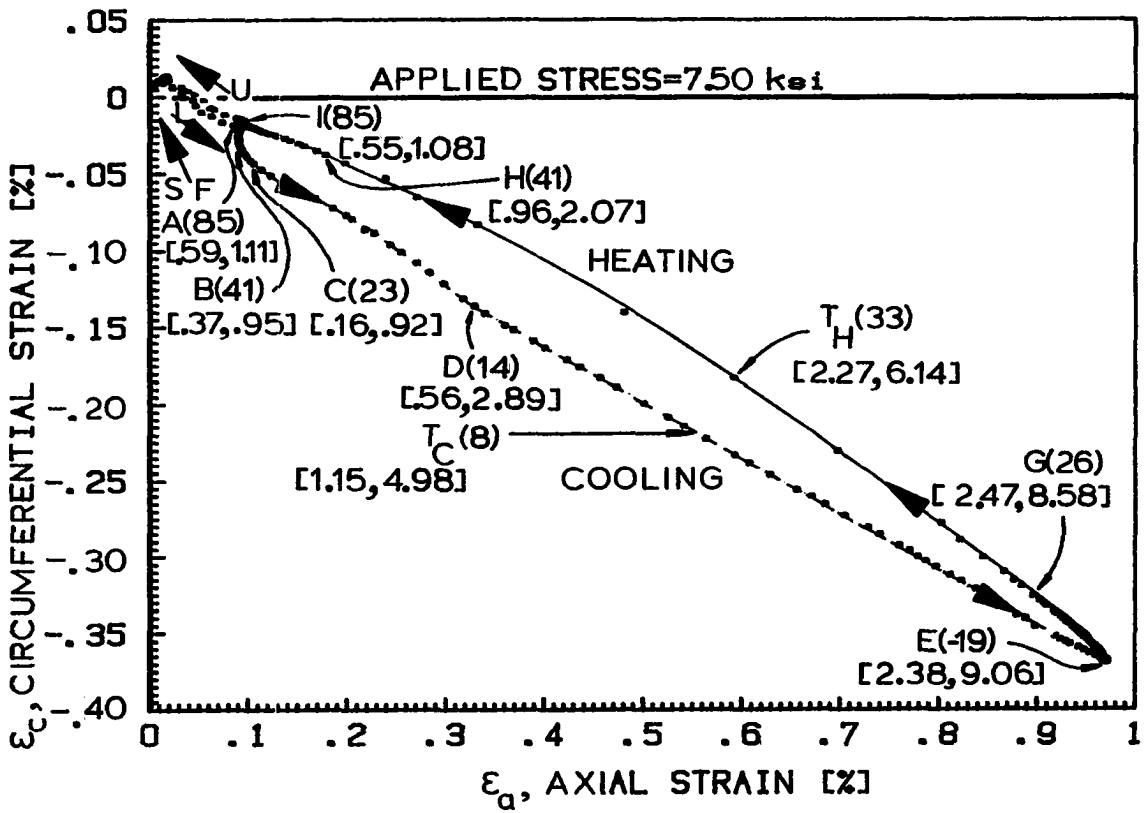


Figure 41. The functional relationship between the circumferential and axial strains for sample P8-1 under 7.50 ksi applied stress. Cooling from A to E (dashed line), and heating from E to I (full line)

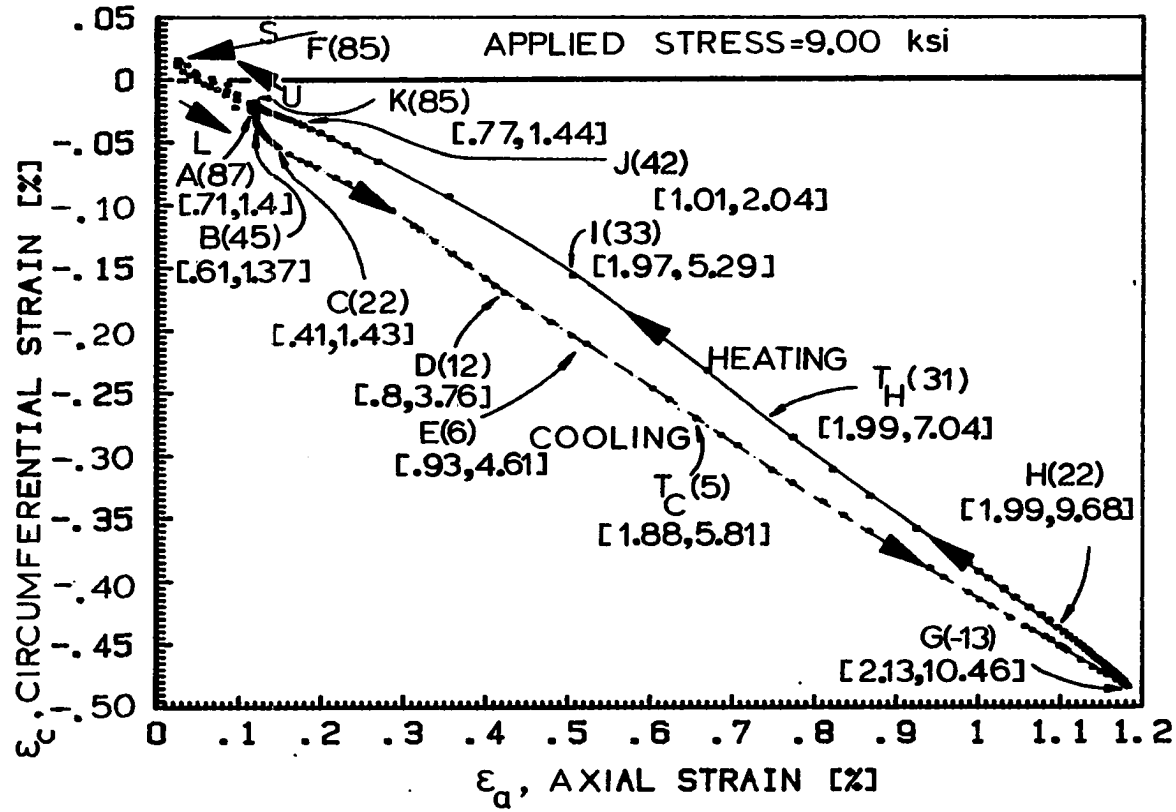


Figure 42. The functional relationship between the circumferential and axial strains for sample P8-1 under 9.32 ksi applied stress. Cooling from A to G (dashed line), and heating from G to K (full line)

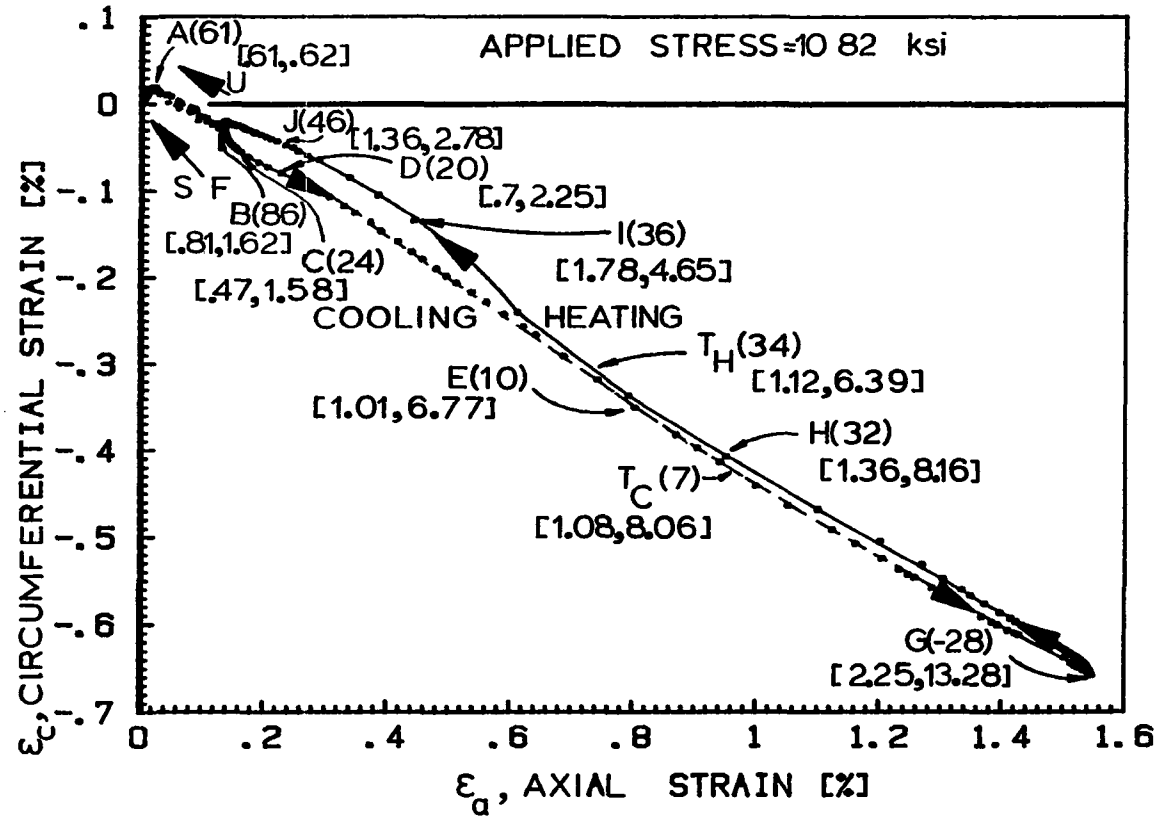


Figure 43. The functional relationship between the circumferential and axial strains for sample P8-1 under 10.82 ksi applied stress. Cooling from B to G (dashed line), and heating from G to near J (full line)

unmeasured transverse strain is the average of ε_a and ε_c. Also, $\varepsilon_a=0$ and $\varepsilon_c=0$ are shown by heavy lines.

A common feature of these figures is their irregular and more or less elongated "banana" shape. The generated loops occupy mainly the fourth and to a lesser extent, the first quadrants of the ε_a-ε_c plane. Also, the change, say, from cooling to heating results in cusplike behavior of the loops, as the cusp at point G in Fig. 37.

With regard to ε_c-ε_a loops, such as in Fig. 37, the direction of the tangent line at any point indicates the state of stress, as is illustrated schematically in Fig. 38. The tangent line at point A on the ε_c-ε_a loop makes polar angle θ with the positive ε axis. The direction $\overrightarrow{OC}$, drawn from the origin also with polar angle θ, lies in the second quandrant for which the stress state is uniaxial compression. In general, polar angles lying in the first, second, third, and fourth quadrants indicate approximate stress states of biaxial tension, uniaxial compression, biaxial compression, and uniaxial tension, respectively, due to the applied stress and to internal thermal, transformation, and residual stresses. In this manner, a state of stress can be attached to each point on the loop. For the zero-load run (Fig. 40), the cooling half-cycle extends from S to near H (dashed curve) and the heating half-cycle from near H to F (full curve). The stress state is biaxial compression from F to C, which occurs mainly because of thermal contraction. At C (21°C), the start of transformation becomes apparent, producing transformation strains that are influenced by the prior history or training of

the sample, i.e., by the prior runs under uniaxial applied stress. This means that the variants forming initially, i.e., for points near and below C, are oriented so as to produce axial expansion and circumferential contraction, appropriate to uniaxial tension. It is also noted by the $\Delta V/V_0$ values (in units of 10^{-3}) in the square brackets that the volume is decreasing at this portion of the curve (approximately from B to D). This decrease in volume is further documented below from measurements on sample P7-1 and may be due to R phase formation. The state of uniaxial extension continues from C(21°C) to E(9°C) in Fig. 40, but as E is approached, the variants become more and more random, whereupon the uniaxial tension ends at E. With the variants forming at random in the range from E to H (and any R phase formation completed), the axial and circumferential strain changes become more ad more appropriate to an isotropic volume expansion characteristic of the B2 to distorted B19 monoclinic transformation [80]. This volumetric expansion with ε_a and ε_c both increasing is representative of biaxial tension. At H, the heating half-cycle starts and the biaxial tension continues to J (22°C). At or near J the reverse transformation to B2 starts, which is accompanied by a decrease in ε_a and a continuing increase in ε_c, thus corresponding to uniaxial compression. At K(30°C), the circumferential direction starts to contract (in addition to the axial contraction) which corresponds to biaxial compression. Starting at K, we have biaxial contraction where the random variants revert to the B2 structure causing a decrease in both ε_a and ε_c due mainly to the isotropic volume contraction.

From M to N, the favorable variants revert to austenite causing a decrease in ε_a and an increase in ε_c, corresponding to uniaxial compression.

In Figs. 37, 39, 41, 42, and 43, the ε_c-ε_a loops are given for runs under applied uniaxial stresses of 3.00, 6.00, 7.50, 9.00, and 10.82 ksi, respectively. From these figures, a clear decrease in the loops' width with increasing applied stress is observed. A similar analysis of the loops for these loaded runs to that given above for the zero-load run of Fig. 40 indicates that the state of stress, internal and applied, corresponds mostly to uniaxial tension in the cooling half-cycle (for example, from A at 85°C to E at -19°C in Fig. 41), and to uniaxial compression in the heating half-cycle (from E at -19°C to I at 85°C in Fig. 41).

The decrease in the width of the ε_c-ε_a loops with increasing applied uniaxial stress and the different states of stresses exhibited by loops of zero-load but trained material compared to loops of runs with applied stress is explainable on the basis of the multiplicity of variants for the martensite. In the absence of load and training, the activation of variants on cooling is random, and on the average, the phase transformation strains are isotropic. Hence, a biaxial stress distribution results. On heating, the process is reversed, and the biaxial stresses fade out as the phase transformation progresses. In the case of uniaxially loaded runs (as illustrated in Fig. 41), however, the applied load tends to polarize the stress distribution in the axial direction, and, hence, to enhance the relative contribution of those variants that

are the most able to contribute to stress release in the applied stress direction. As a consequence, the stress distribution is uniaxial tension on cooling and uniaxial compression on heating. In the intermediate zero-load runs on previously trained material (Fig. 40), with the onset of MPT on cooling, the favorable variants will be activated due to the training effect. They contribute a uniaxial tension to the state of total stresses even in the absence of an applied stress. As the phase transformation progresses, the average variants take over and the state of stress becomes more and more that of biaxial tension, as was observed.

Another approach can be taken by recognizing that in an elastic material under uniaxial applied stress the axial and the circumferential strains are in phase. They are coupled by a single constant, Poisson's ratio. Their graphical representation for unidirectionally changing stress is a straight line.

The loops generated in the ε_c-ε_a plane in our case are not straight lines, but rather irregular banana shapes. However, as the applied load increases, the loops straighten and become narrower with increasing applied stress (Fig. 43). This effect is not due to different scaling of the axes for loaded and zero-load runs, as presented in Figs. 44 and 45, respectively. In these figures, the same loops were drawn as in Figs. 39 and 40, respectively, but the axes were normalized to $\varepsilon_a(\max)$ and $\varepsilon_c(\max)$ for each case. The result remains the same. The loop for the loaded run is straight compared to the loop for the zero-loaded run.

Each point on the normalized loops of Figs. 44 and 45 is defined in polar coordinates by the distance of the point from the origin and by the angle between the segment connecting the origin with the point of interest and the horizontal axis, ε_a=0. The range of values of this angle, for the two cases, emphasizes the already stated result, namely, in case of loaded runs the two strains, ε_c and ε_a, are very much in phase compared to those of the zero-load runs.

This concept of phase angle can be developed further. Although the shape of the loops are far from being an ellipse, major and minor axes can be defined in the following way. During thermal cycling the material transforms from HTP to LTP and back to HTP. Thus, the LTP and HTP points are two end points of the forward and reverse transformation, respectively. The segment connecting them is defined as the major axis, a, and the segment which is cut out from the bisectional normal to the major axis and the loop, as the minor axis, b. The ratio of these axes, e=b/a, can be used as a figure of merit to indicate roughly the average phase shift between the axial and the circumferential strains during thermal cycling as a function of applied stress. The result is shown in Fig. 46, confirming that increasing applied uniaxial load tends to bring the axial and circumferential strains in phase with one another with increasing effectiveness.

If the two strains were in phase, then the loop generated in the ε_c-ε_a plane would degenerate to the major axis. Also, the tangent of the angle between the major axis and the ε_a=0 line should give, in this

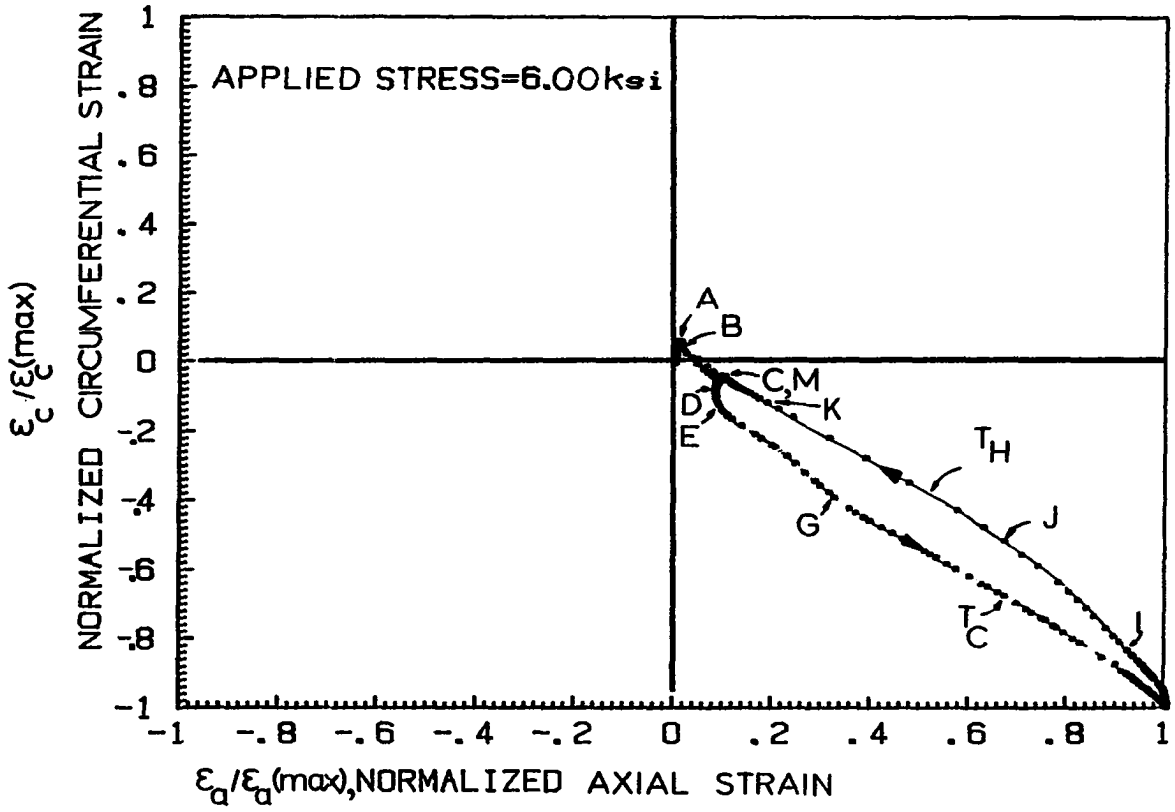


Figure 44. The same plot of Fig. 39, but normalized to maximum axial, $\varepsilon_a(max)$, and circumferential, $\varepsilon_c(max)$, strains. Cooling from C (dashed line), and heating from near I to M (full line)

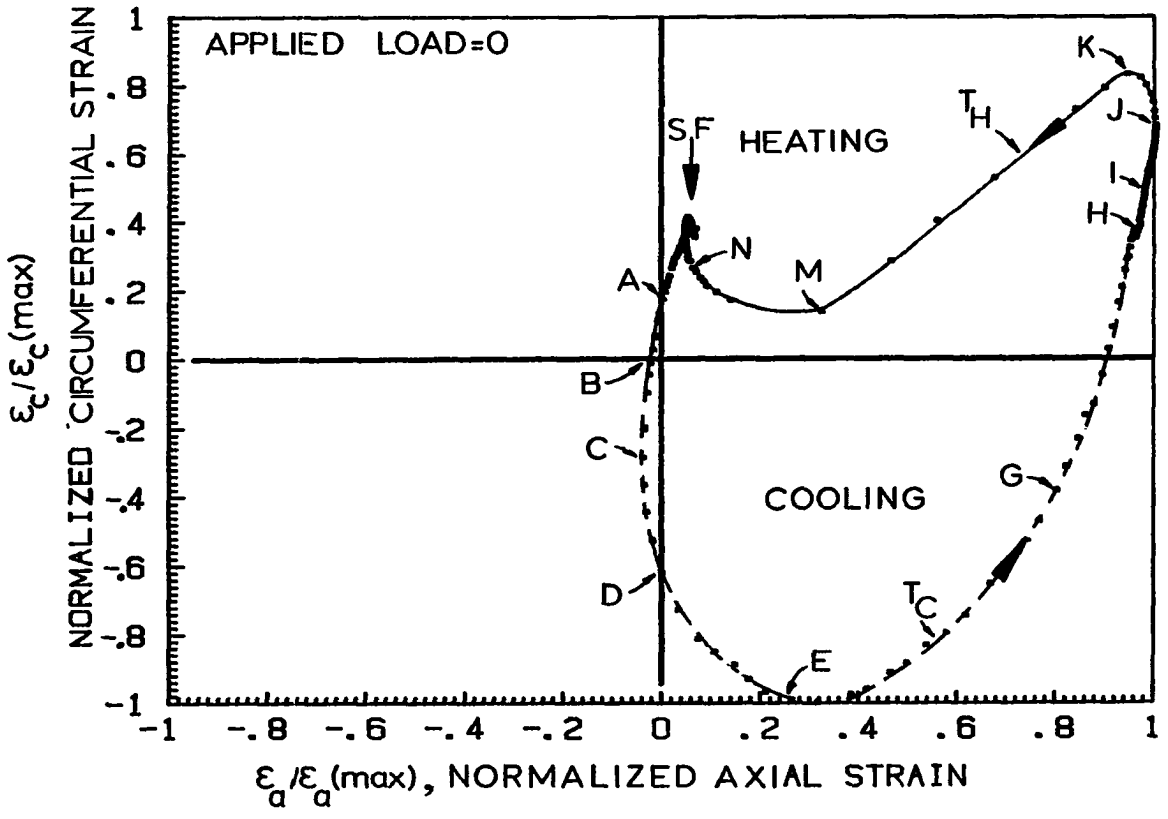


Figure 45. The same plot as of Fig. 40, but normalized to maximum axial, $\varepsilon_a(max)$, and circumferential, $\varepsilon_c(max)$, strains. Cooling from S to near H (dashed line), and heating from near H to F (full line)

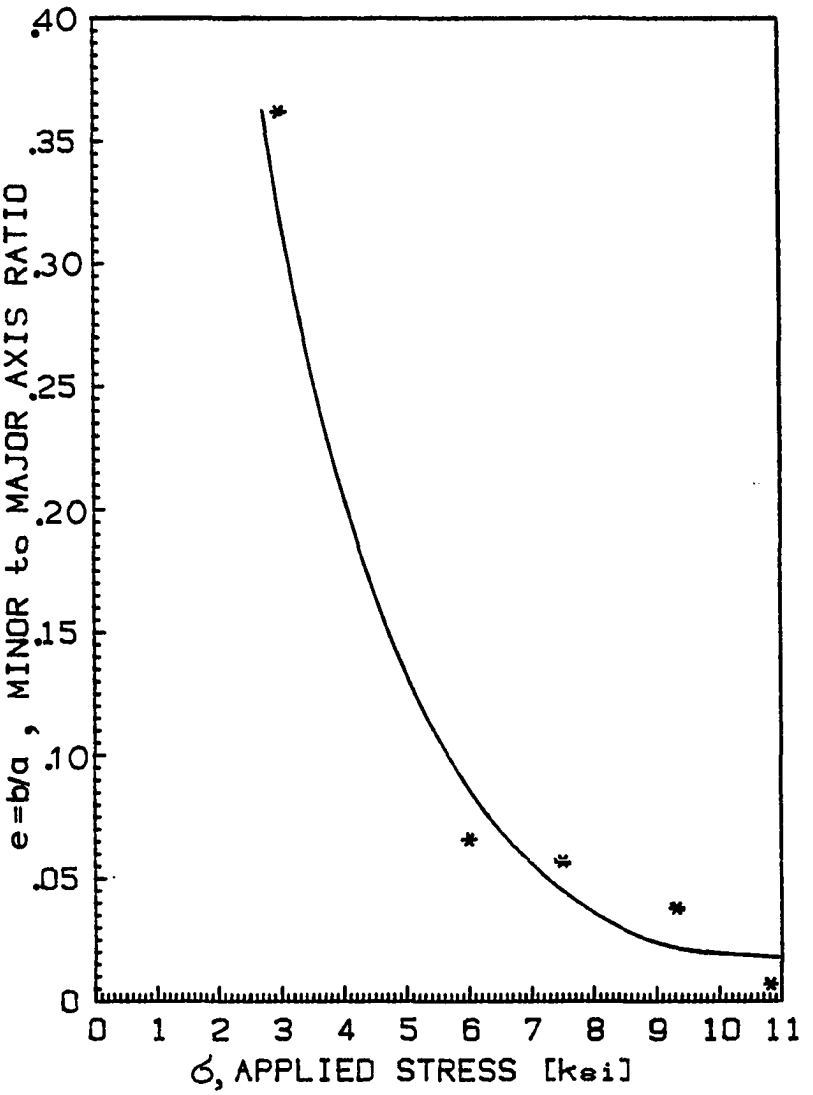


Figure 46. Ratio of minor, b, to major, a, axes of the transformation loops, e=b/a, as a function of applied stress

case, the negative of the Poisson's ratio, $\tan\psi=-\nu$. As shown in Fig. 47, $-\tan\psi$ approaches the value of ~0.45 for applied stress of 10.82 ksi, as ν steadily increases from the value of ~0.27 at 3.00 ksi applied stress.

In the HTP, the structure is pure B2; hence, it should behave as an elastic body under loading below the yield stress. Strain values taken during loading and unloading of sample P8-1 in the high temperature phase at 85°C were used to plot Figs. 48 and 49, from which by least squares linear fitting the Young's modulus, Y, and the Poisson's ratio, ν, were calculated and found to be 10.1×10^6 psi and 0.44, respectively, in agreement with data from the literature [29]. Hence, the applied stress not only causes the two strains to be more in phase, but to be in phase along a line that has a negative slope corresponding to the HTP Poisson's ratio.

3.1.4 The dependence of transformation strain on applied stress

An important aspect of the martensitic phase transformation is the dependence of the strain-temperature behavior on applied stress. Also, the effect of the thermomechanical history (training) is of interest. To investigate these effects, axial and circumferential strains were measured as a function of temperature cycling through the transformation with various constant loads applied to the sample. The results for loaded runs are shown in Figs. 50 and 51 and those for zero-load are shown in Fig. 52.

Let us define the MPT strain (on cooling) as the maximum magnitude of strain during the cooling part of the cycle where the strain is zeroed

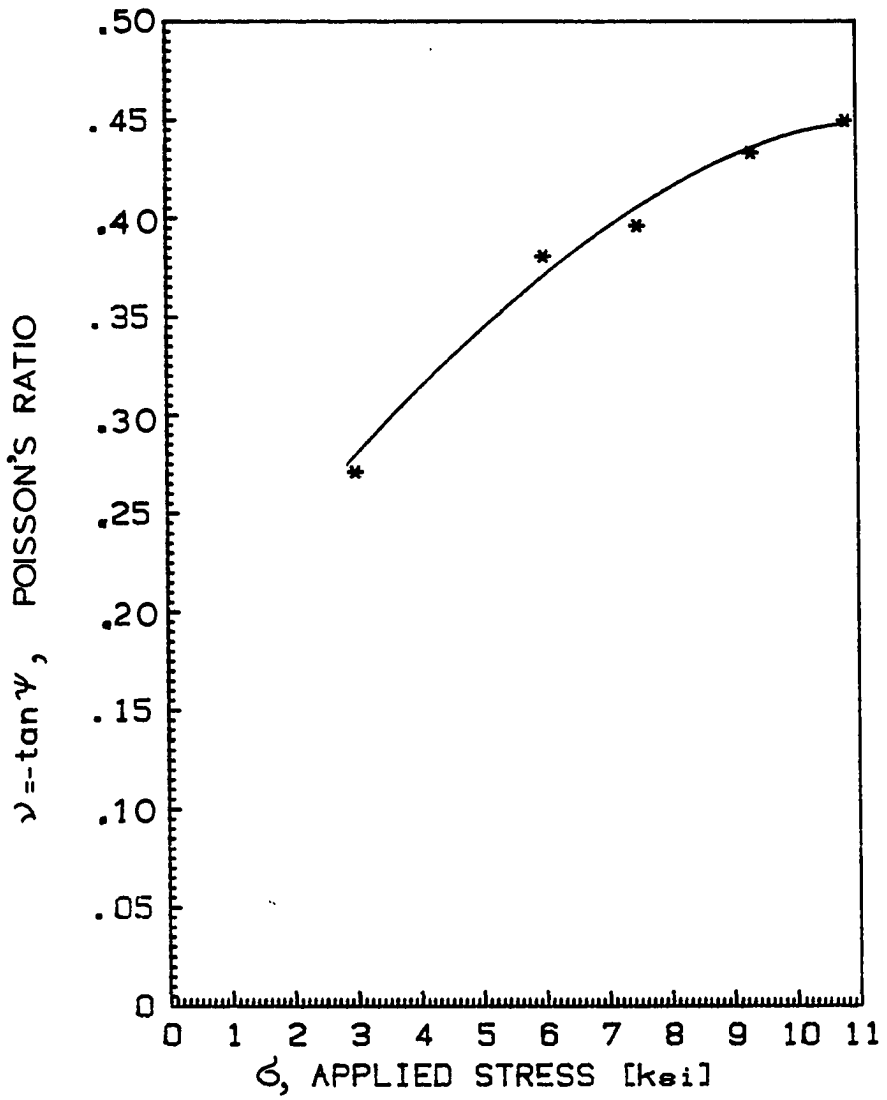


Figure 47. The negative of the tangent of the inclination angle, Ψ, of the major axis with respect to the line ε_a=0 verus applied stress. ν=-tan Ψ, where ν is the Poisson's ratio

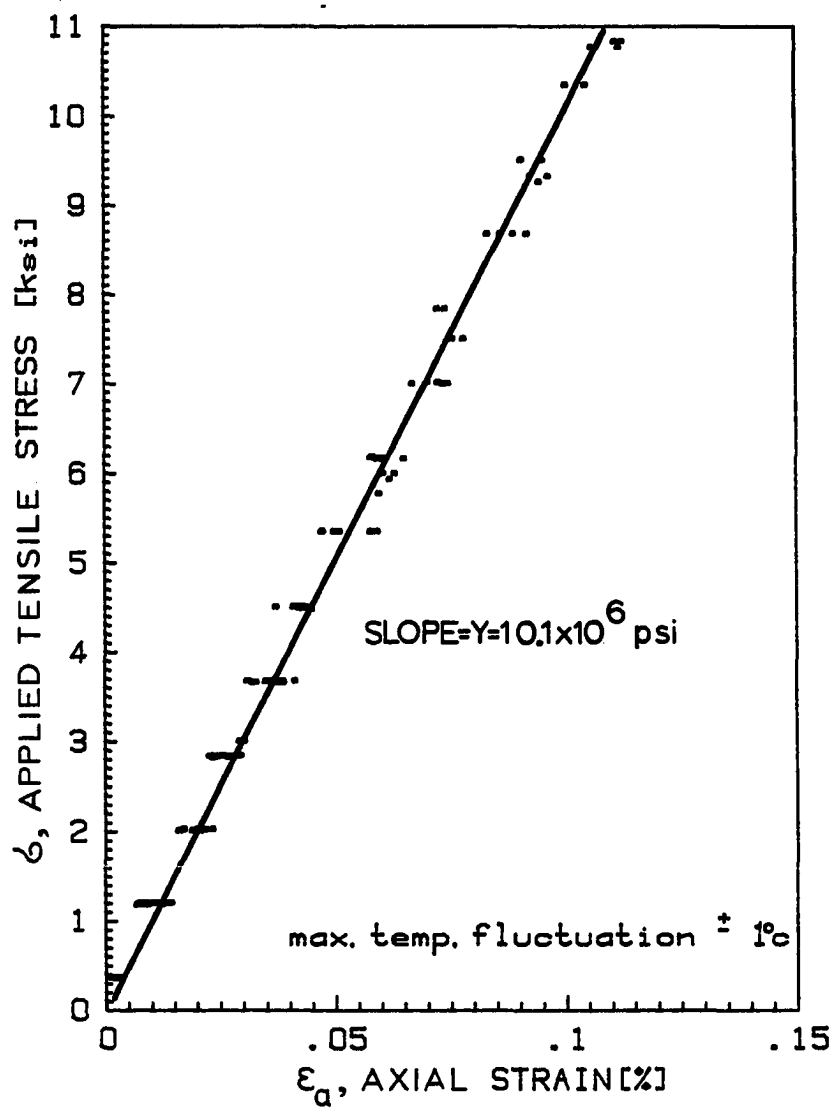


Figure 48. Least-square fit correlation line between applied stress and axial strain from data points of loading and unloading sample P8-1 in the HTP, at 85°C

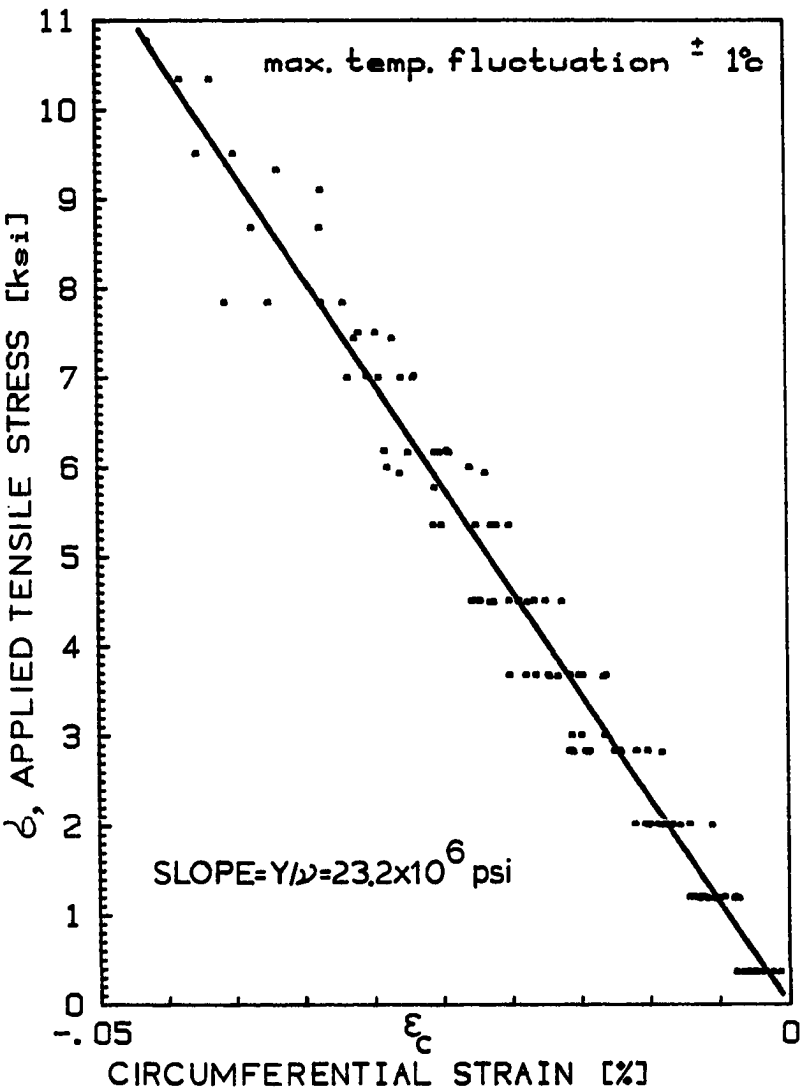


Figure 49. Least-square fit correlation line between applied stress and circumferential strain from data points of loading and unloading sample P8-1 in the HTP, at 85°C

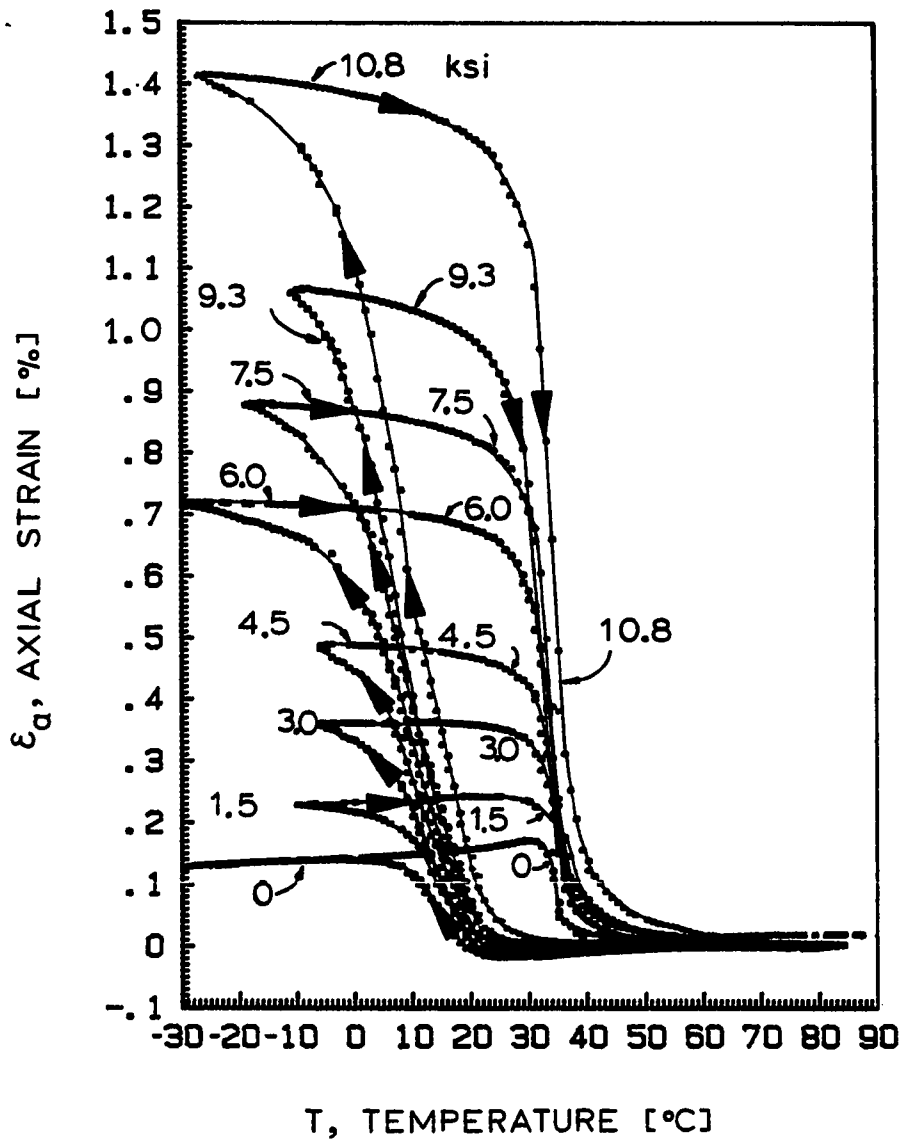


Figure 50. Summarizing figure of axial strain versus sample temperature as a function of applied stress for sample P8-1

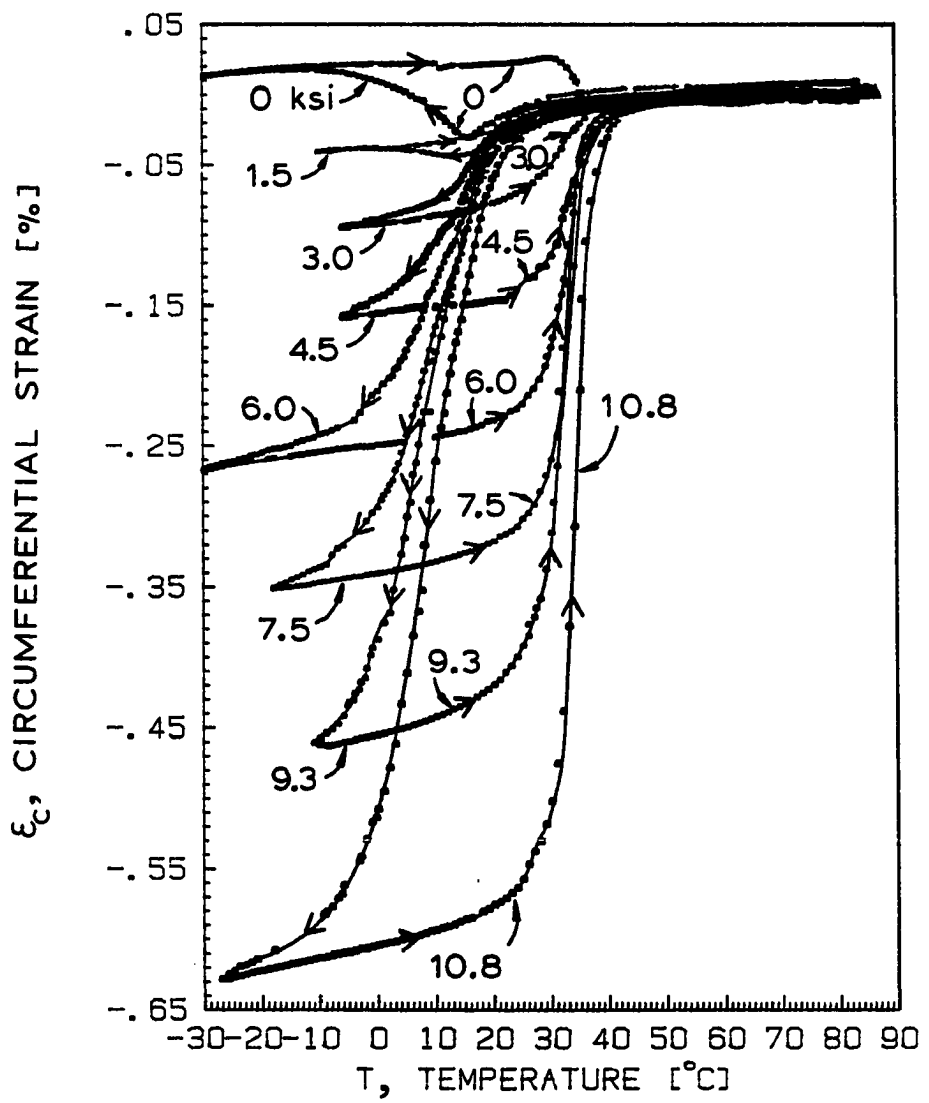


Figure 51. Summarizing figure of circumferential strain versus sample temperature as a function of applied stress for P8-1

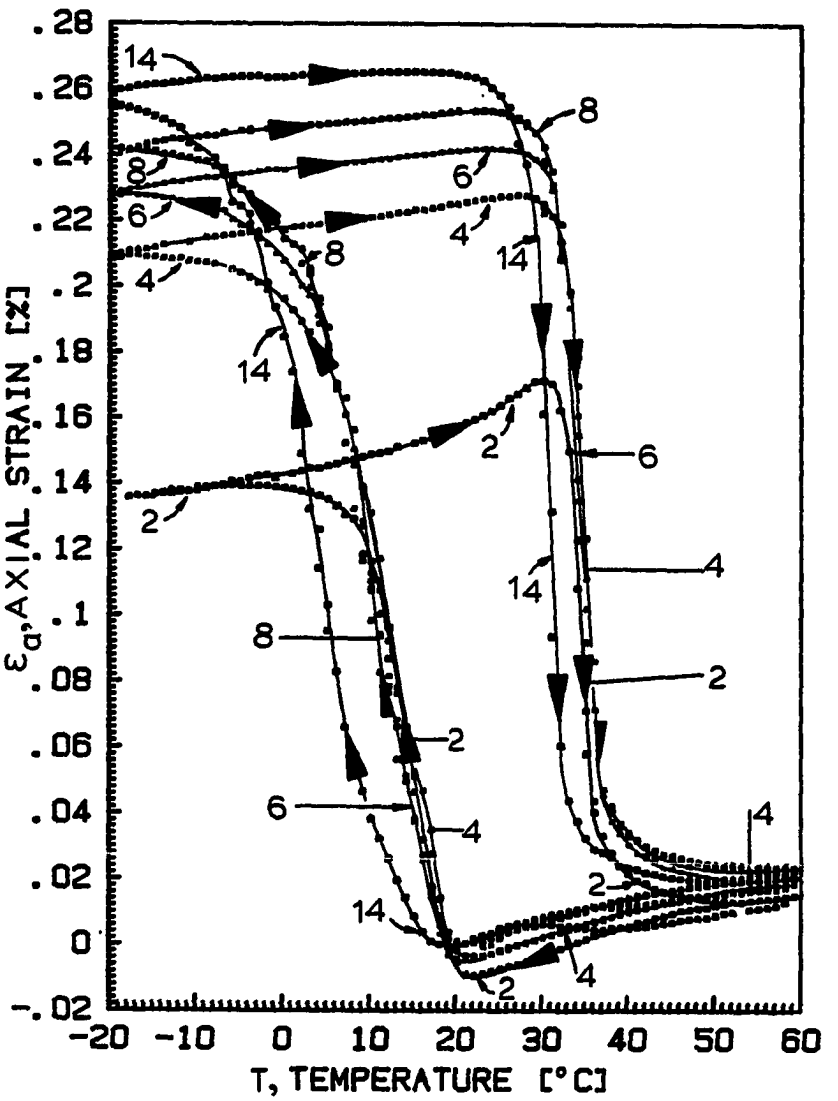


Figure 52. Summarizing plot of axial strain versus sample temperature for zero load runs as a function of increasing run numbers for sample P8-1

at the HTP (at 85°C), as shown in Fig. 53. It is clear from Figs. 50 and 51 that the transformation strain increases with increasing tensile stress, as was observed also by Hsu and Wechsler on cast and swaged samples [79]. This increase occurs both for the axial and the circumferential strains, although in different senses. Upon cooling under an applied uniaxial tensile stress, the sample elongates in the axial direction and contracts in the circumferential direction. The increase in phase transformation strain is, however, not linear in applied stress in the range of investigation. Also, the change is greater for the axial transformation strain than for the circumferential transformation strain. These results are shown in Fig. 54, where the axial and the circumferential MPT strains were plotted versus applied stess.

Two factors may contribute to the observations shown in Figs. 50 and 51 that the transformation strains increased in magnitude with increasing applied stress from run to run in the sequence of runs performed: (1) the applied stress itself, and (2) the effect of training, i.e., the fact that the runs were conducted serially so that the results of each run may have affected by having experienced the previous run or runs. The evidence points to (1) as being a more important factor than (2). Firstly, the runs illustrated in Figs. 50 and 51 were not conducted one immediately after the other. Instead, as indicated in Table 2, these loaded runs were separated by a run at zero load. The results of these in-between runs at zero load are shown in Fig. 52. Secondly, the magnitudes of the increases in transformation strain for each second run as the applied

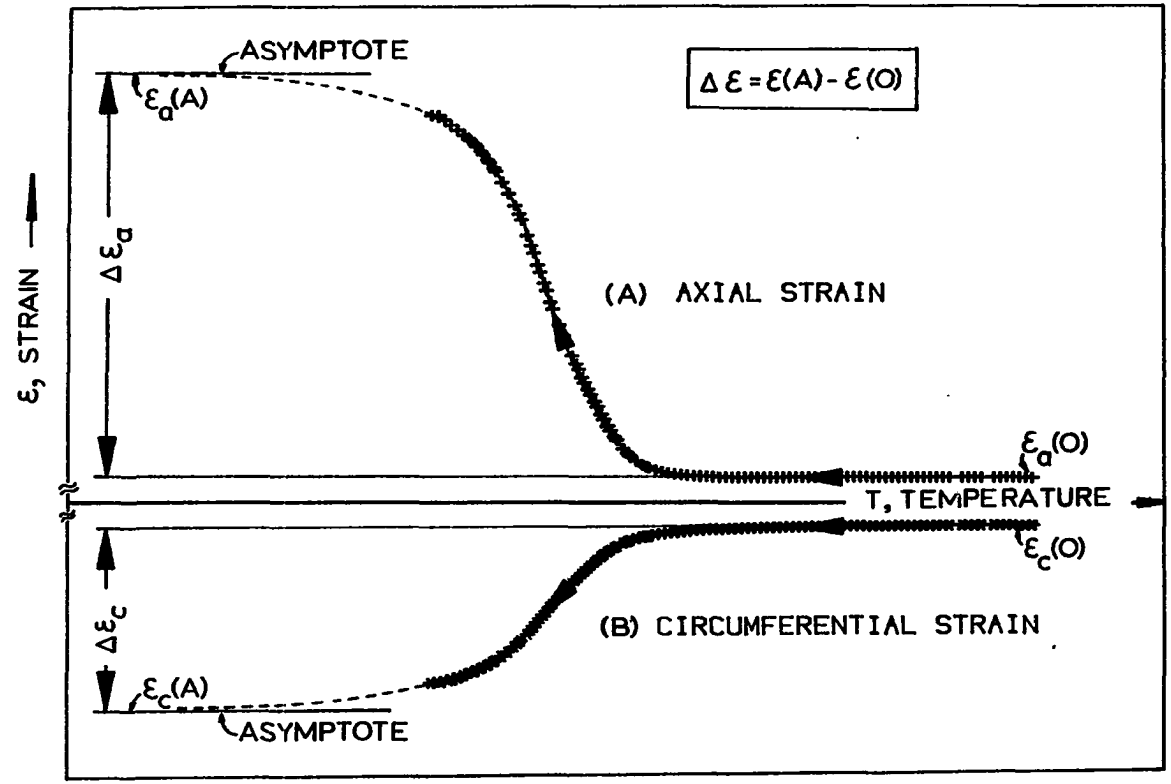


Figure 53. Schematic strain-temperature curves on cooling. (A) axial strain; (B) circumferential strain. Illustration of determination of transformation strains on cooling

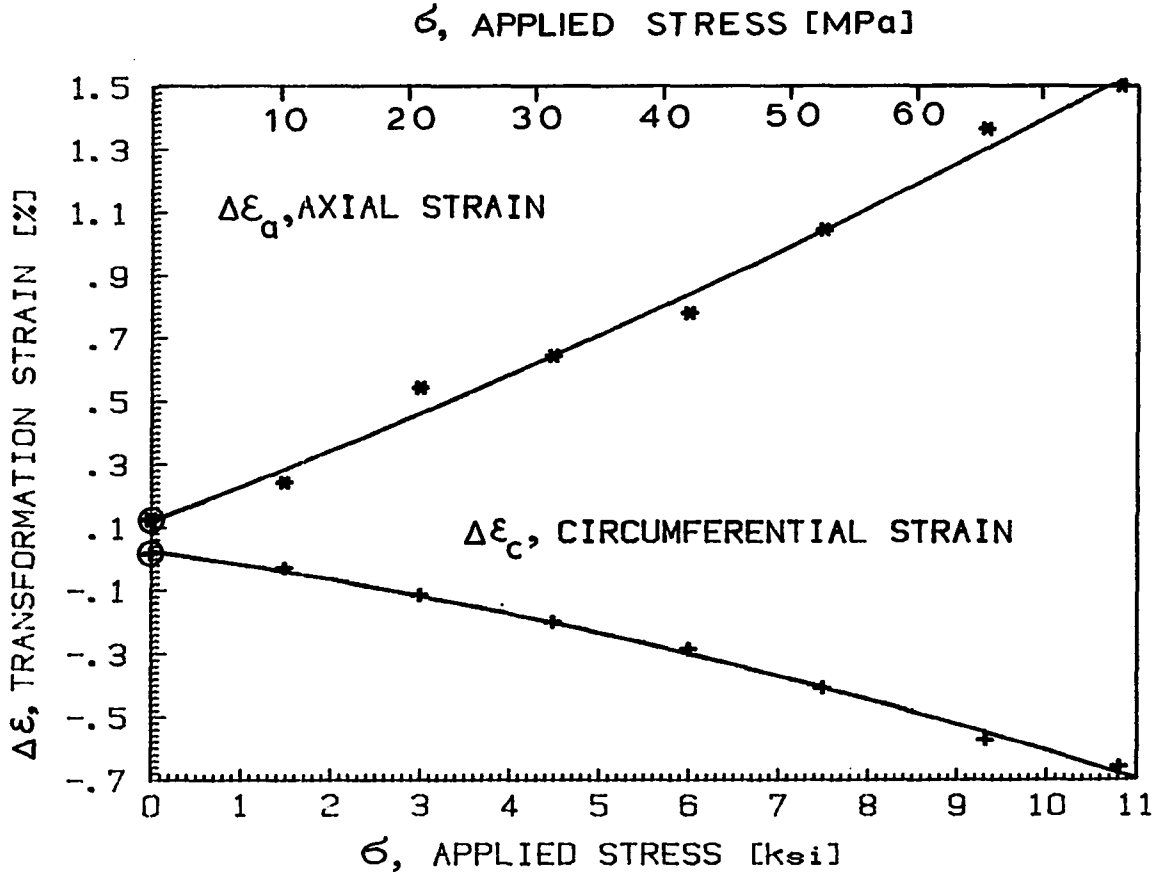


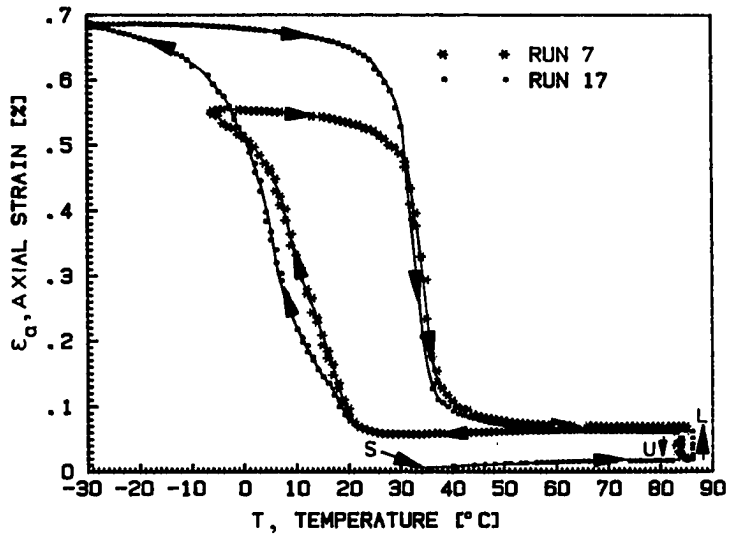
Figure 54. Axial and circumferential transformation strains versus applied stress for sample P8-1

stress was increased was much greater than the increase in transformation strain for repeated runs at the same stress. Thus, the axial transformation strain increases from 0.12% at zero load to 1.5% at 10.82 ksi (Fig. 54), whereas they increased from 0.12% for the first zero-load run to only 0.26% for the seventh zero-load run (Run 14, Fig. 52). Furthermore, the results of run 7 at 4.56 ksi may be compared with the results of run 17 at the same stress. Fig. 55 shows that the transformation strains were not greatly different (for the axial strains, about 0.5% for run 7 and 0.65% for run 17). What is concluded from the above discussion is that the effect of stress during cycling predominates, but there is, nevertheless, a clear lesser effect of the training. From these results and the application of the preferred activation of variants model (PAM) (see Appendix C), it is suggested that the variants that are activated at a previous high load will be more probably activated at a lower or zero load than other available variants. The meaning of the above is that the sample develops long range memory during loaded runs, namely it not only regenerates its original shape upon heating, but also "remembers" to some extent the path it followed during previous, highly loaded run, on cooling. This is a kind of two way-memory.

The increasing response of transformation strain to increasing applied stress may be regarded in two ways:

(a) Phase transformation in absence of applied and residual stresses should occur more or less isotropically in strains, especially in randomly oriented fine-grain samples. Thus, the transformation strains

(a)



(b)

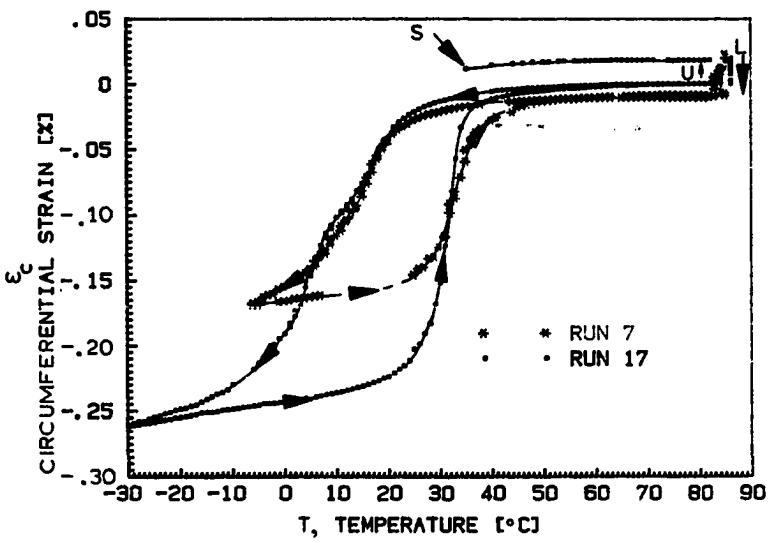


Figure 55. Comparison of curves of (a) axial and (b) circumferential strains for Runs 7 and 17 of sample P8-1 under applied stress of 4.5 ksi

should and do change, in the same sense along the axial and circumferential directions of the sample (expansion on cooling). However, when stress is applied, favorable variants are preferred statistically, as a trade-off between the zero-load untrained MPT tendency and the stress induced situation. A break-even point seems to occur at about 2 ksi and at higher stresses variants with preferred orientation dominate (see twisted circumferential hysteresis loop at 1.5 ksi applied stress in Fig. 21(b)). Since the preferred activation of variants is probabilistic, and the probability increases with increasing applied stress, the magnitude of the strain hysteresis loop slope is affected, as shown in Fig. 50. Hence, for higher applied stress, smaller temperature difference provides the same amount of change in strain as for lower stress levels. This fact contributes partially to the increased transformation strain response with increasing applied stress.

(b) Another phenomenon which contributes, springs from the nucleation temperature of phase transformation. This temperature is higher for the preferred variants (the subcooling range is shorter) than for the average, as preferred variants are partially stress promoted, and hence M_S is higher for higher applied stresses. This stress dependence of M_S is also a corollary from thermodynamic considerations as can be shown by the use of the Clausius-Clapeyron equation [88]. For the same reason, variants that are strongly not preferred have lower nucleation temperatures than the average, pushing M_F down to lower temperatures (see Fig. 36). Hence, the temperature range over which forward phase

transformation takes place increases with increasing applied stress. So even if the slopes of the strain temperature hysteresis loops were equal for all cases of loaded runs independent of the applied stress (in contradiction to (a)), the transformation strain would increase with increasing applied stress.

The combined effect of applied stress on transformation strain as was argued above in (a) and (b) can be expressed mathematically in the following way. Equation (4) in case of a finite volume ΔV of martensite formation can be written as:

$$\Delta G = -\Delta g_v \Delta V + \Delta g_a \Delta A + \left(\sum_{jk}\sum \int \sigma_{jk}^{(i)} d\varepsilon_{jk}\right) \Delta V - \int \sigma^{(a)} d\varepsilon_a \Delta V \qquad (5)$$

where ΔA is the corresponding increase in surface area. Let p be the probability of transformation of the small volume of austenite. Then, this probability behaves as

$$p \propto \exp\left(-\frac{\Delta G}{kT}\right) \qquad (6)$$

As it was discussed previously, ΔG depends among other variables, on the applied stress, the temperature, and indirectly on the variant involved in the transformation. Hence, for a given variant, 1, at temperature T the probability of transformation behaves as

$$p_1 \propto \exp\left(-\frac{\Delta G_1}{kT}\right)$$

and similarly for another variant, 2, as

$$p_2 \propto \exp\left(-\frac{\Delta G_2}{kT}\right)$$

It follows that for these two variants, the relative probability to be produced upon transformation is

$$\frac{p_1}{p_2} = \exp\left(\frac{\Delta G_2 - \Delta G_1}{kT}\right) \qquad (7)$$

Now, let us assume that the first variant is favored, i.e., its contribution to ΔG is highly negative through the fourth term on the right hand side of equation (5), and the second variant is not a favored one. Then, assuming that the third term in equation (5) does not change much for the two cases compared to the large negative value of the fourth term for the favored variant case,

$$\Delta G_1 < \Delta G_2 \quad \text{and} \quad p_1 >> p_2$$

and more so the higher the applied stress.

Hence, the probability that MPT will occur through the activation of the favored variant is overwhelmingly larger than that for the less favored variant. But, as was discussed earlier, favored variants contribute to the transformation strain mainly in the direction of applied stress (hence, their effect as described above in (a)), and they cause increase in the temperature range over which transformation takes place (as is described above in (b)). Hence, the observed increase in the transformation strain with increasing applied stress is explained as a consequence of preferred activation of variants.

The "anomalies" in the differential strain curves versus temperature are also quite understandable. When the load is applied, the sample is in the HTP. As the load is relatively low, the applied stress induces

only elastic deformation of the sample, namely a small elongation along its axis and the related Poisson's contraction in the circumferential direction. At the start of the MPT, the first martensite is formed in the austenite matrix. Its transformation strain must be accommodated. Without applied load, the martensite variant selection is random to ensure most economic accommodation from the point of view of energy balance. However, with applied tensile load the variant that will contribute to the maximum elongation of the sample is preferred, as in this case the load, and hence the center of gravity of the system will be lowered. The potential energy decreases and hence, the total energy. The preferred variant is the one for which the MPT extension along the sample axis tends to relieve the axial stress. Thus, these variants will be activated preferentially, causing elongation along the sample axis, and contraction in the circumferential direction, as was observed.

As the MTP proceeds, the transformation strains due to the preferred variant build up, and it becomes energetically favorable for a second variant to form whose transformation strains tend to cancel those due to the first variant. The axial strain ε_a for the second variant is smaller than for the first variant and therefore the decrease in ΔG due to the fourth term in the right-hand side of equation (5) is reduced. But, this more than offset by the reduction in ΔG due to the reduction in the third term in equation (5) due to the cancellation of transformation strains. From this scenario it is clear that the preferred variants will be trained more, cycle after cycle, than the others, and hence, when a

zero load run follows a highly loaded one, the most trained variants will have a larger statistical probability to operate than the others. But when they do so at zero applied load the decrease in ΔG due to the fourth term in equation (5) is absent, and the energy balance is upset. Hence, the other variants take over.

For a zero-load cycle following a high-load cycle, at the beginning of transformation the differential strain curve behaves as if the cycle was loaded. Depending on training, i.e., on the magnitude of the load on the previous run or runs, the differential curve, after the initial transient behaves as for a normal zero-load cycle (Fig. 32).

The above PAM model differs from other models [36, 63, 89, and 90] by stating that the major effect of applied stress on phase transformation strain is due to the preferred activation of a martensite variant and not due to the reorientation of the martensitic structure.

3.1.5 Practical consequences of transformation strain dependence on applied stress

The interest in transformation strain dependence on applied stress is not merely theoretical, but also practical. The work output of any solid-state low-temperature heat engine (one of the possible applications of shape memory alloys) depends both on phase transformation strain and on applied load. In the particular case of the uniaxially loaded tubular sample P8-1, work can be expressed as the product of the applied stress, axial phase transformation strain, and volume of gage section. In Fig. 56, the work values are given for the case of a tube dead-weight loaded

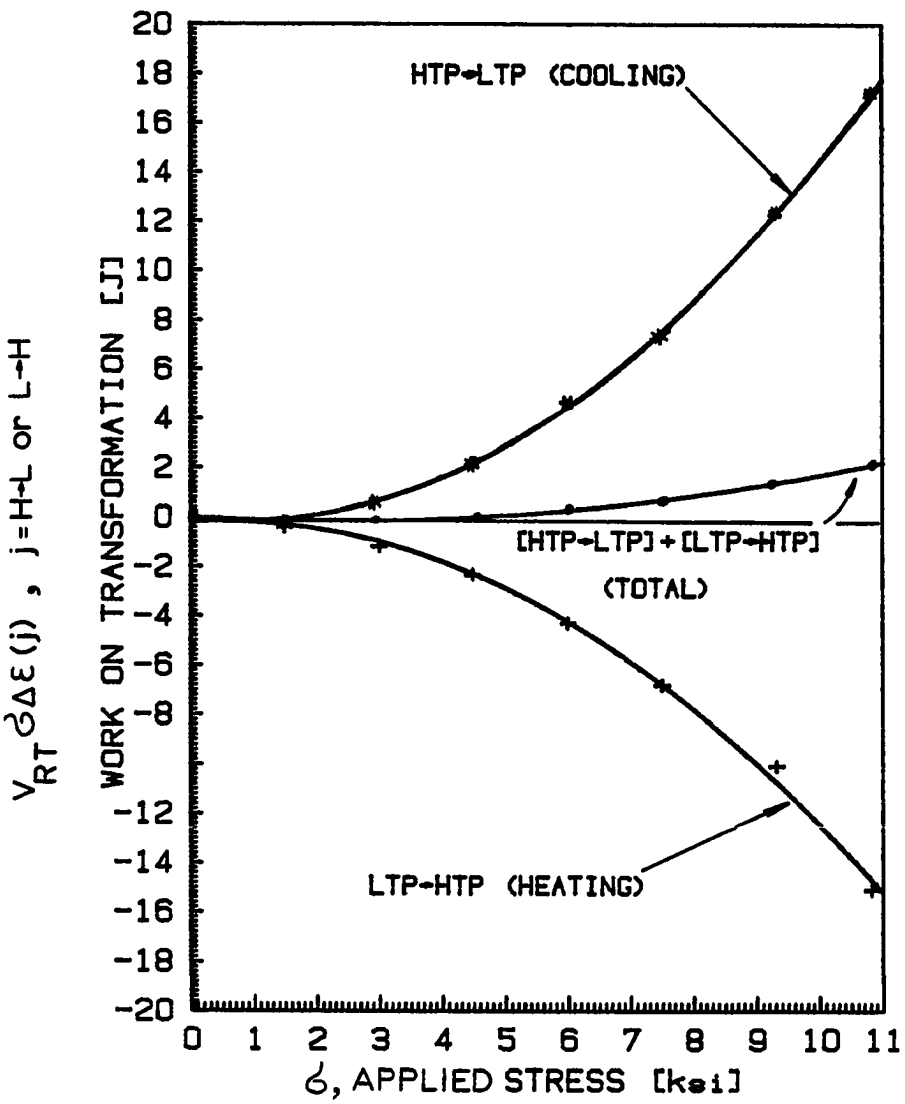


Figure 56. Work on transformation of the gage length part of sample P8-1

in tension during cooling and heating half-cycles. The tube expands on cooling and work is done on the tube by the external dead-weight load (positive work). On heating, the tube contracts and it does work in lifting the weight (negative work). The sum of these is the total work done on the tube over the entire cycle. At the highest stress of 10.82 ksi, the work done by the tube during transformation on heating is about 15 joules and the work done on the tube during transformation on cooling is about 18 joules. Over the entire cycle, a net amount of work of about 3 joules is done on the tube. In a practical shape memory heat engine, the load would be removed or decreased for the cooling half-cycle, so that net work would be done by the shape-memory metal on its surroundings [91].

3.2 Thermal Cycling of Annealed Sample PO-4 Under Applied Tensile Stress

Following vacuum anneal at 600°C for 2 hours and furnace cooling, sensors were attached to the sample surface as shown in Fig. 14. An additional two element 90° strain gage rosette was mounted on the sample and its pickups connected in a half bridge manner to the strain indicator device. Thus, the temperature (T), axial strain (ε_a), circumferential strain (ε_c), strain in 45° to the sample axis (ε_{45}), and shear strain (γ_{zc}) were measured simultaneously during the thermal cycle. From the measured data, the major (ε_p) and minor (ε_q) principal strains were calculated by:

$$\varepsilon_{p,q} = \frac{\varepsilon_a+\varepsilon_c}{2} \pm \frac{1}{\sqrt{2}} \sqrt{(\varepsilon_c-\varepsilon_{45})^2 + (\varepsilon_{45}-\varepsilon_a)^2} \qquad (8)$$

and the maximum shear strain by:

$$\gamma_{max} = \sqrt{2} \sqrt{(\varepsilon_c-\varepsilon_{45})^2 + (\varepsilon_{45}-\varepsilon_a)^2} \qquad (9)$$

From the geometry of the sample, and previous experience with sample P8-1, it was assumed that the general behavior of ε_a and ε_p is similar, and also for ε_c and ε_q. Thus, for each run ε_a and ε_p are plotted together, and similarly for ε_c and ε_q. The applied tensile stress during this experiment had the values: 0, 0.5, 1.0, 1.5, 3.0, 6.0, and 9.0 ksi.

The results for the tensile strains (ε_a, ε_p and ε_c, ε_q) are given in Fig. 57 for the first zero load cycle. The hysteresis loops are of small magnitude and their overall change is in the same direction, i.e., expansion on cooling and contraction on heating, as already was observed for the as fabricated P8-1 sample. The measured strains are not zero (although they are small) and the shapes of the hysteresis loops for ε_a, ε_{45}, ε_p, ε_c and e_q, are similar, as is expected since the sample was in the annealed condition and was transformed under zero applied load.

The results for cycling under high tensile stress (not shown) are also as anticipated based on results for the as-fabricated sample P8-1.

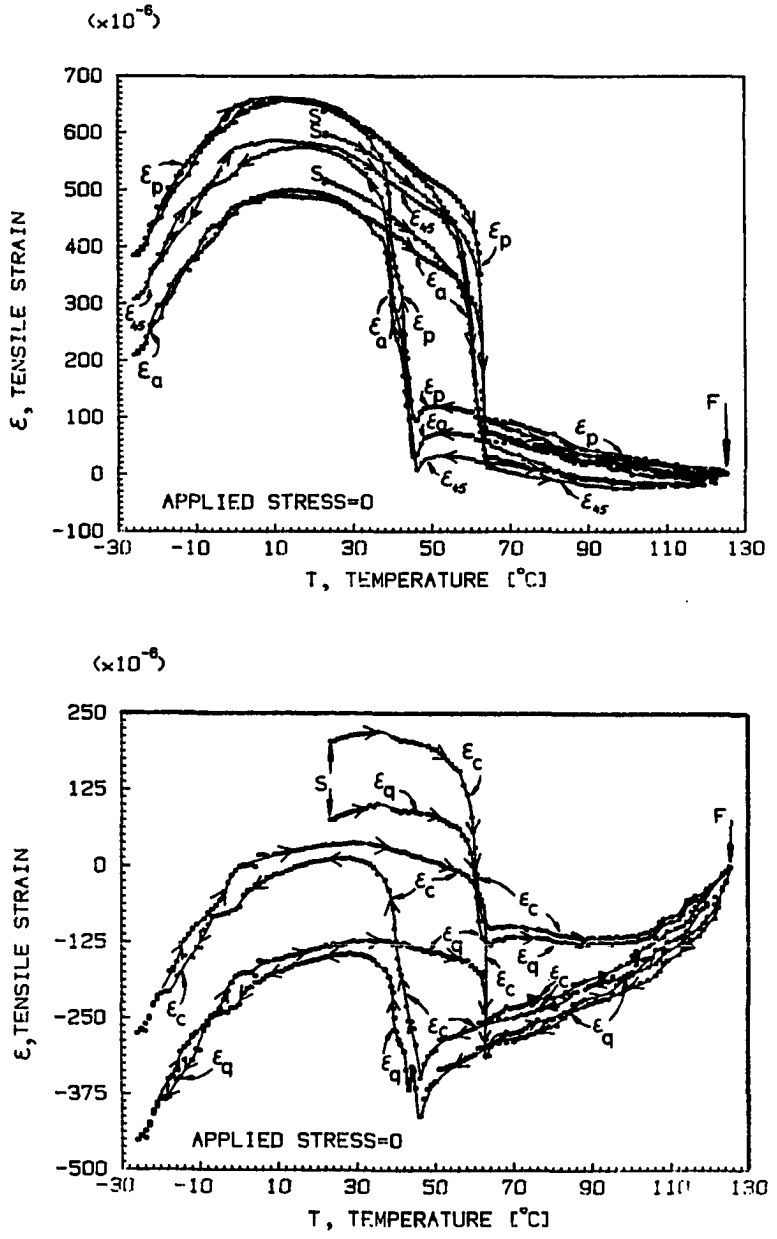


Figure 57. Strain-temperature curves of (a) experimentally measured axial and 45° inclined strains and calculated major principal strain, and (b) experimentally measured circumferential strain and calculated minor principal strain of annealed sample PO-4 on the first thermal cycling under zero applied load

ε_p and ε_q achieve high values, their general shapes are similar, but their directions of change on MPT are opposite; i.e., ε_p increasing on cooling, giving a positive transformation strain, while ε_q decreasing on cooling, contributing a negative (contraction) transformation strain. Moreover, ε_a and ε_p become equal, and also ε_c and ε_q. In addition, ε_{45} becomes small compared to any other measured or calculated strains in the other directions.

The most striking results were obtained during thermal cycling under low applied stresses. It was pointed out for the as-fabricated sample P8-1 that strain hysteresis loops are twisted at low applied stress (see, for example, Fig. 21(b)). The effect was related to a transition range in which the applied stress is not high enough to overcome the tendency of the applied stress-free MPT to exhibit a small expansion on cooling in all directions (random activation of variants). However, in case of sample P8-1, only the circumferential strain hysteresis loop was found to be twisted and not the axial. As only the circumferential strain hysteresis loop direction was changed with increasing applied stress, from assumptions of continuous behavior it was quite satisfying to observe the twisted loops. The axial strain hysteresis loop direction was unaltered by the increasing applied stress, and hence intuitively there was no need for twisting.

If the PAM is correct, the above-mentioned transition range must exist in all directions, and not only in the circumferential. Of course, twisting should not happen necessarily at the same applied load for

strains in all directions. As the applied tensile stress increases, more and more variants are activated whose directions of maximum extension (Fig. C2, Appendix C) are aligned near the axial direction. However, a given distortion ellipsoid can be rotated about its major axis such that the smallest axis will line up in a different direction, thus reducing the contraction effect in the circumferential direction, but still contributing the same elongation along the sample axis. Hence, we expect that the transition range for the axial direction will occur at lower applied stresses than for the circumferential direction. Thus, Fig. 58(a) shows the twisted loop at 0.5 ksi for the axial strain, but Fig. 58(f) shows it slightly at 1.5 ksi for the circumferential strain and more clearly at 2.0 ksi (not shown). From these values it is clear, why a twisted loop was not observed for the axial strain of the as-fabricated P8-1 sample. There the first applied stress was 1.5 ksi, which is above the transition range for the axial direction. For circumferential direction, however, 1.5 ksi is near transition range for twisted loops, and hence, the twisting for this direction was observed (Fig. 21(b)).

The transformation strains for the axial and circumferential directions as function of applied stress are shown in Fig. 59(a) and (b), respectively. Their general behavior is very much the same as for the as-fabricated P8-1 sample (Fig. 54), except for a small irregularity in the curves at low applied stress (near 0.5 ksi). As already was explained, this regime is the transition regime. Above the transition regime, all the transformation strains behave regularly, i.e., $\Delta\varepsilon_p$,

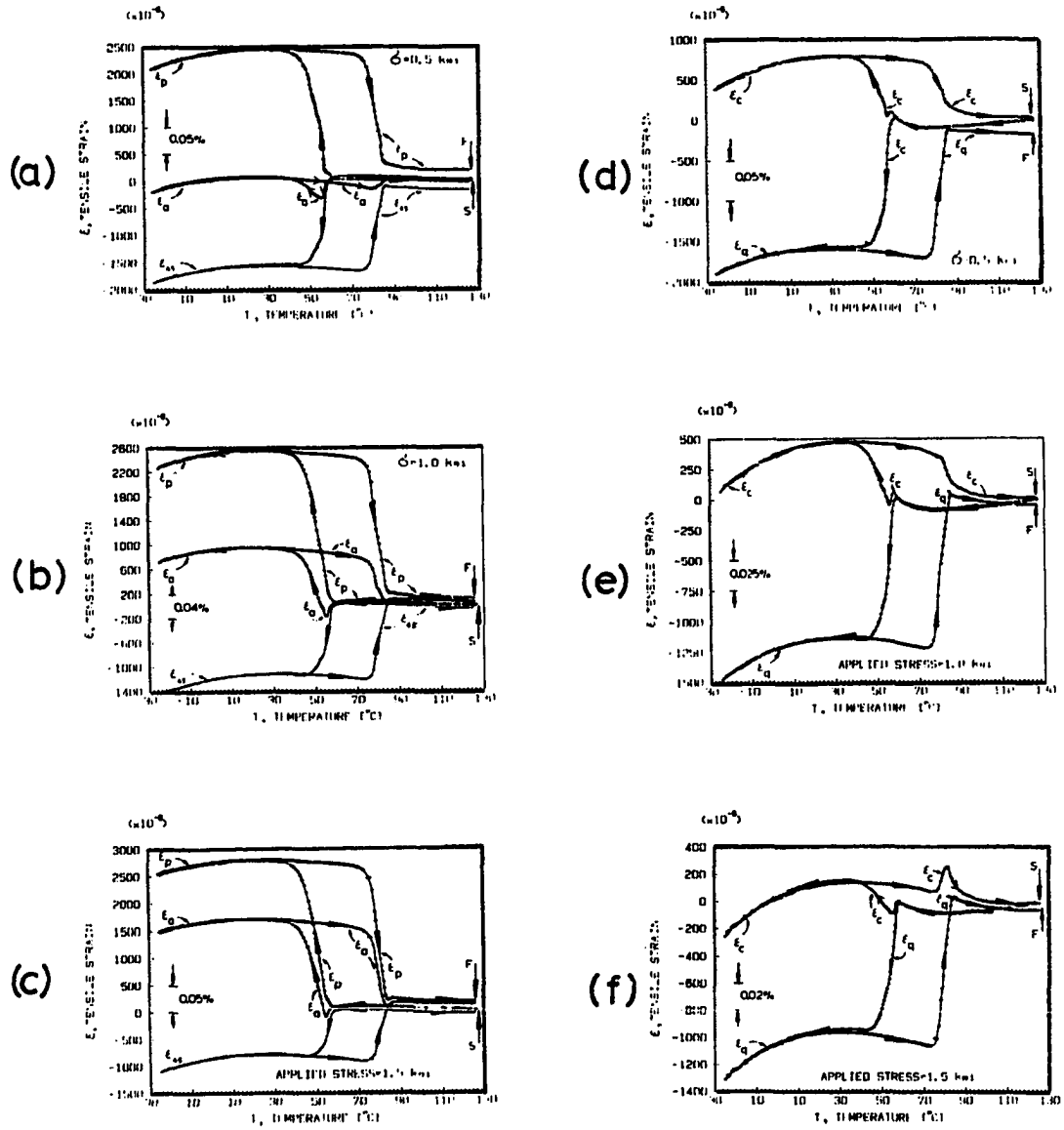
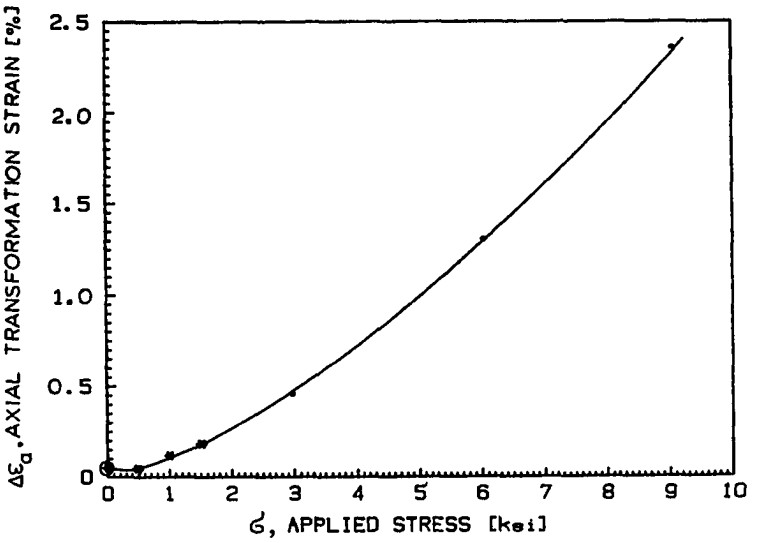


Figure 58. Strain hysteresis loops for various applied stresses. (a), (b), and (c): ε_p, ε_a, and ε_{45} for stresses of 0.5, 1.0, and 1.5 ksi, respectively. (d), (e), and (f): ε_q and ε_c for stresses of 0.5, 1.0, and 1.5 ksi, respectively

(a)



(b)

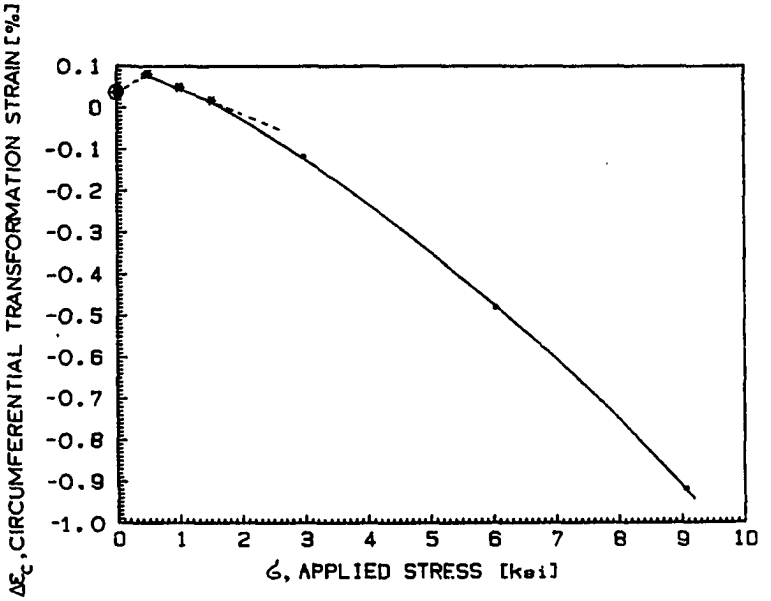


Figure 59. Strain-stress curves of (a) axial and (b) circumferential strain in annealed sample PO-4 as a function of applied uniaxial tensile stess

$\Delta\varepsilon_a$, and $\Delta\varepsilon_{45}$ increase with increasing applied stress, and the sample expands in these directions upon forward phase transformation (cooling), while $\Delta\varepsilon_c$, and $\Delta\varepsilon_q$, increase in magnitude with increasing applied stress, but the sample contracts in these directions (Fig. 60).

All the curves are concave in Fig. 60, away from the zero horizontal line ($\Delta\varepsilon$=0), suggesting that the effect of applied stress on the phase transformation strain increases with increasing stress. The increase of the magnitude of $\Delta\varepsilon$ with increasing σ_a must be limited by another factor. Such a factor is the build-up of internal strain energy as the amount of a single variant increases in a given region, i.e., the increase in the third term on the right-hand side of equation (5). When the internal strain energy becomes too high, another variant must form that produces a smaller strain in a particular direction.

For uniaxial tension, at constant load the extreme case is where the third term in equation (5) is sufficiently small that it can be ignored. Then, the distortion ellipsoids are all aligned with their major axis near the tensile direction, thus on cooling contributing the largest possible elongation in this direction. This upper limit was calculated in Appendix B, using the phenomenological theory of WLR for the case of NiTi. According to the result, the theoretical limit of the MPT strain in the tensile direction ($\Delta\varepsilon_a^{th}$) is about 5.4%. Hence, it is anticipated that the dependence of $\Delta\varepsilon_a$ with applied stress will change direction at high applied stresses, aiming toward the theoretical limit in an asymptotic manner.

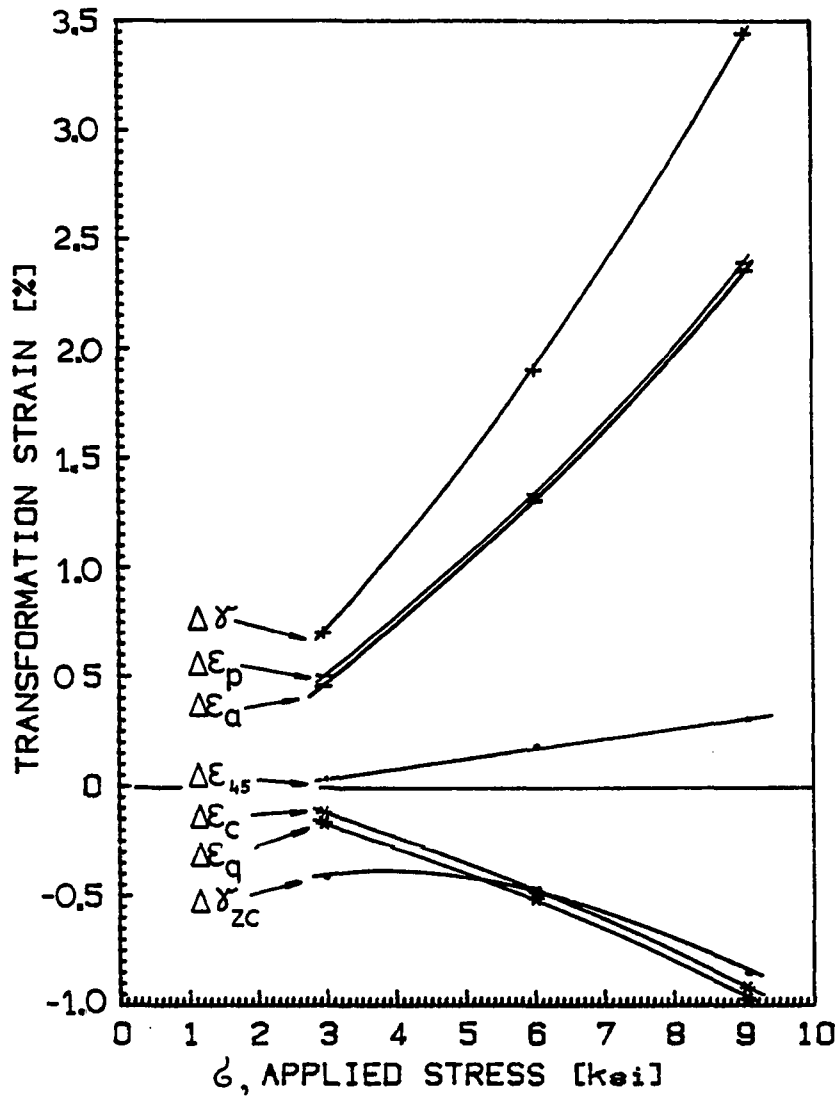


Figure 60. Transformation strains on cooling for annealed sample P0-4 as a function of applied uniaxial tensile stress for stresses of 3 ksi and higher

Using the assumptions of Mohamed (equation 6 in ref. [92]), we can write:

$$\sigma_a^{th} = \sigma_y^{HTP}(1-\exp(-N\Delta\varepsilon_a^{th})) \qquad (10)$$

where: σ_a^{th} is the theoretical limit of applied uniaxial tensile stress at which all the deformation elipsoids are aligned with their largest axis along the tensile direction, σ_y^{HTP} is the yield stress of material at the HTP, N is the number of crystallographically equivalent variants, and $\Delta\varepsilon_a^{th}$ is the MPT strain in the tensile direction corresponding to σ_a^{th}. Substituting the numerical values:

$$\sigma_y^{HTP} = 45 \text{ ksi [93]}$$

$$N = 24$$

$$\Delta\varepsilon_a^{th} = 0.054$$

get σ_a^{th} = 33 ksi, as the applied stress at which all the deformation ellipsoids theoretically should be aligned.

3.3 Thermal Cycling of Annealed Sample P8-1 Under Applied Tensile and Compressive Stresses

The annealed P8-1 sample was thermal cycled in tension between zero and 7.8 ksi applied stress and in compression between near zero (-36 psi) to -7.6 ksi, according to the following pattern:

(a) Tension with increasing load from cycle to cycle: 0, 0.5, 1.0, 1.6, 3.2, 6.1 and 7.8 ksi.

(b) Tension with decreasing load from cycle to cycle: in the reverse order to that in (a) down to 34 psi(0^+).

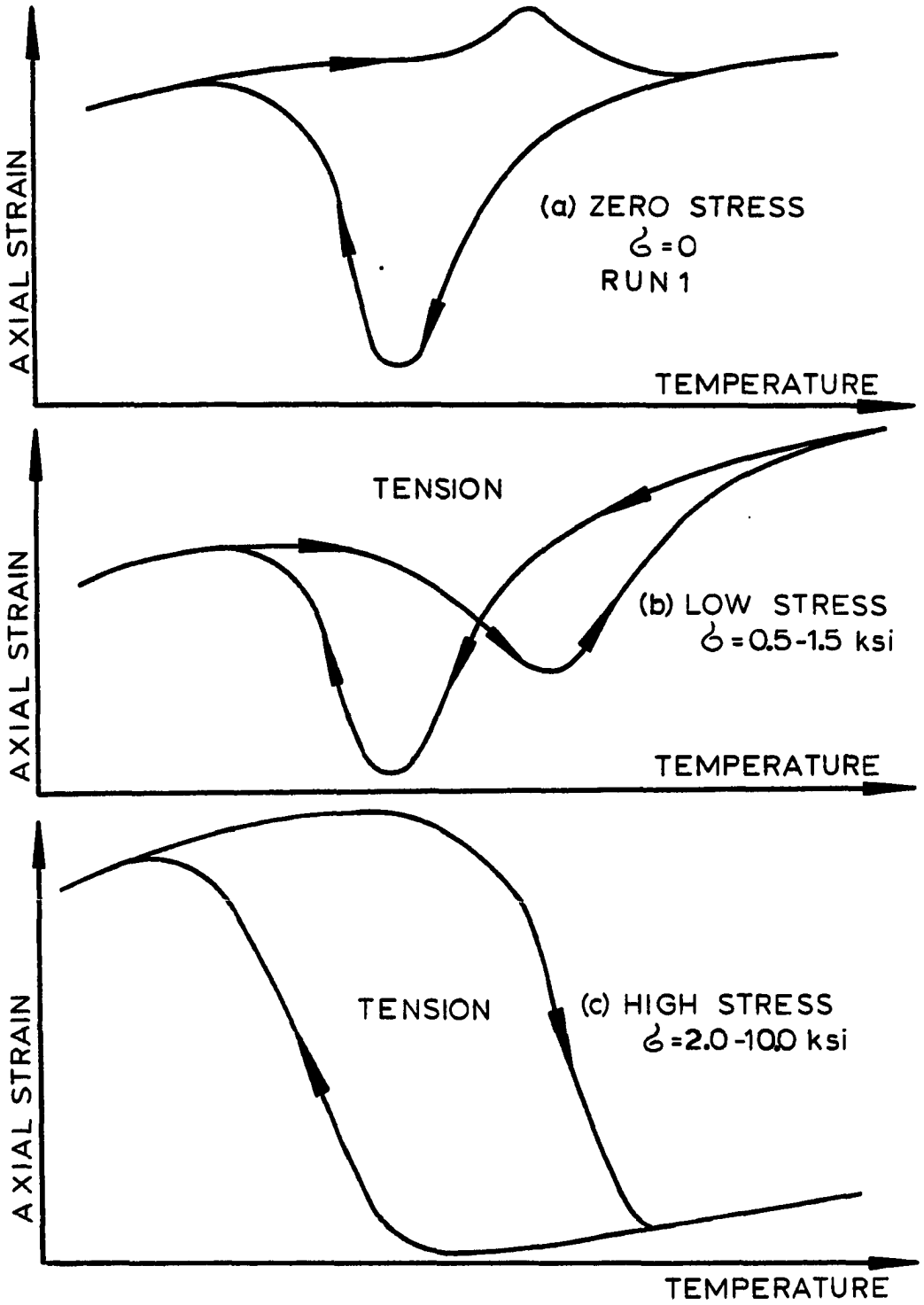
(c) Compression with increasing compressive load from cycle to cycle: 0^-, -0.3, -0.8, -1.3, -2.6, -5.9, and -7.6 ksi.

(d) Compression with decreasing compressive load from cycle to cycle: in the reverse order to that in (c) down to -36 psi (0^-).

When the forward and reverse transformation generated strain hysteresis loops were investigated for the various cases given in (a)-(d), it was found that a kind of "fingerprint" of the thermomechanical history of the sample is reflected in the strain hysteresis loops. Five different types of hysteresis loops were found in the case of axial strain (and similarly for the circumferential strain) as shown schematically in Fig. 61. According to these prototype hysteresis loops, a small nearly zero MPT strain exists for an annealed sample (Fig. 61(a)), a transition range both on tension (Fig. 61(b)) and compression (Fig. 61(d)) manifested by loop twisting, and the normal behavior at high stress level, expansion on cooling for tension (Fig. 61(c)), and contraction on cooling for compression (Fig. 61(e)).

The normal behavior (Figs. 61(c) and (e)) is the same as was observed by Wasilewski [67] and explained (also not on the PAM basis) by Delaey et al. [63]. The experimental results are displayed in Fig. 62, showing that at high tensile stresses ε_a and ε_p are superimposed, and similarly ε_c and ε_q, while at high compressive stresses ε_a and ε_q are superimposed, and similarly ε_c and ε_p. Thus, in compression the

Figure 61. Schematic axial strains versus temperature for: (a) zero applied stress, (b) applied tensile stress from 0.5 to 1.5 ksi, (c) applied tensile stress from 2.0 to 1.0 ksi, (d) applied compressive stress from 1.0 to 8.0 ksi


AXIAL STRAIN
(a) ZERO STRESS
σ = 0
RUN 1
TEMPERATURE
AXIAL STRAIN
TENSION
(b) LOW STRESS
σ = 0.5-1.5 ksi
TEMPERATURE
AXIAL STRAIN
TENSION
(c) HIGH STRESS
σ = 2.0-10.0 ksi
TEMPERATURE

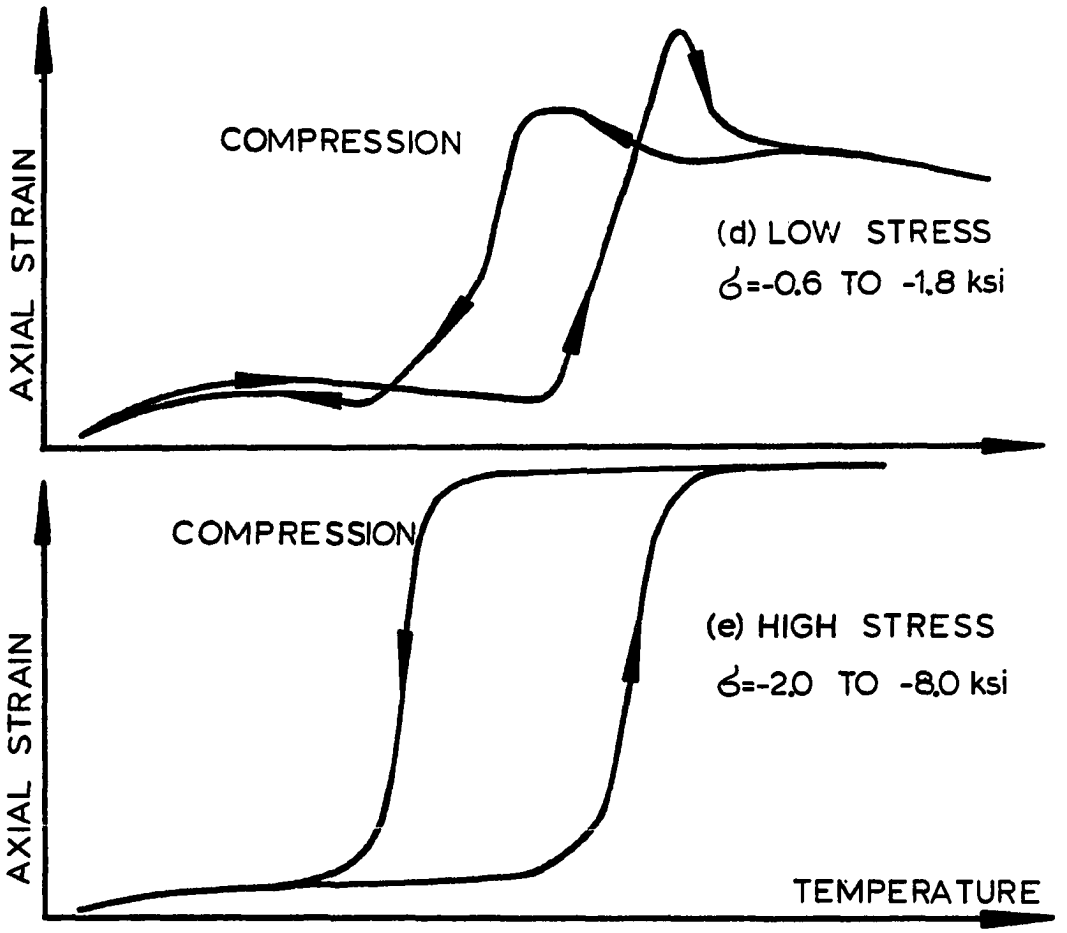


Figure 61 (continued)

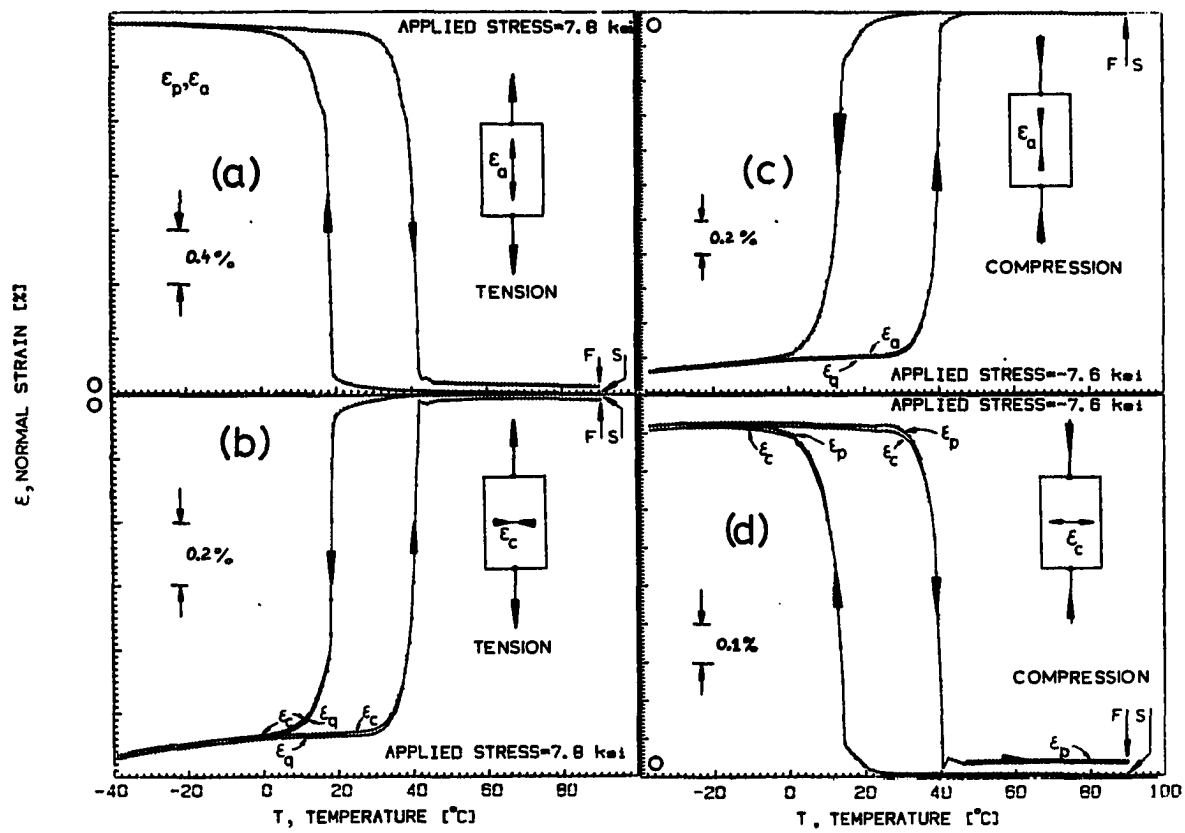


Figure 62. Normal strains versus temperature for sample tube P8-1 for (a) ε_a and ε_p for applied tensile stress 7.8 ksi, (b) ε_c and ε_q for applied tensile stress 7.8 ksi, (c) ε_a and ε_q for applied compressive stress 7.6 ksi, and (d) ε_c and ε_p for applied compressive stress 7.6 ksi. ε_a and ε_p are superimposed in (a); ε_c and ε_q are superimposed in (b), except where indicated; and similarly for (c) and (d)

major principal axis is along the circumferential direction of the sample, while in tension it is along the axial direction. The rotation of the principal strain axis relative to the sample coordinates is energy consuming, as it is done by overcoming and erasing the trained properties. Hence, when a Ni-Ti sample is thermal cycled under bending stresses, unstable regions exist along its neutral line of bending. The contribution of these regions to the total transformation strain is negligible and their existence is parasitic.

How does this rotation of the principal axes of strain take place? In order to investigate this question, the strain hysteresis loops were plotted in Fig. 63, for a small tensile stress (Figs. 63(a) and (b)) followed by a small compressive stress (Figs. 63(c) and (d)). In this figure, ε_a and ε_p are still superimposed both for tension and compression, and similarly for ε_c and ε_q. However, whereas the fitting between the principal axes and the sample axes is perfect in tension, it is less satisfactory in compression, especially at the low temperature part of the hysteresis loops. This continues to be the case for a larger compressive stress (-0.3 ksi applied compressive stress; not shown).

Concentrating on the cooling part of the thermal cycle (and assuming that on heating the process is simply reversed), we found that at -0.8 ksi compressive stress ε_a and ε_p are perfectly superimposed down to about 0°C (and also for ε_c and ε_q), but below this temperature ε_c follows ε_p, and ε_a follows ε_q (Fig. 64(a)). Thus, below 0°C, ε_a is

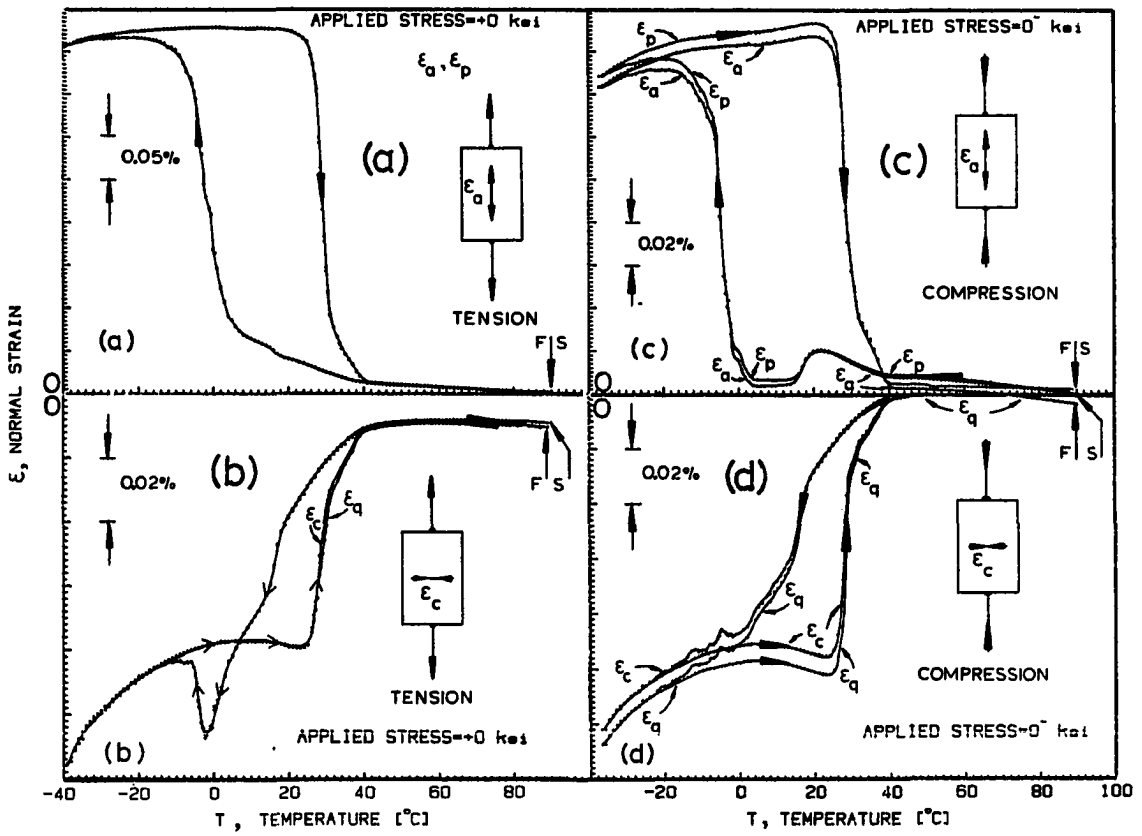


Figure 63. Normal strain versus temperature for sample tube P8-1 for (a) ε_a and ε_p for applied tensile stress near zero (actually 34 psi), (b) ε_c and ε_q for applied tensile stress, near zero (34 psi), (c) ε_a and ε_p for applied compressive stress near zero (36 psi), and (d) ε_c and ε_q for applied compressive stress near zero (36 psi). ε_a and ε_p are superimposed in (a); ε_c and ε_q are superimposed in (b), except where indicated; and similarly for (c) and (d)

approximately the minor principal strain and ε_c the major principal strain. The situation is very similar to that discussed earlier for zero load runs that followed high stress runs (as-fabricated P8-1 sample, section 3.1). The variants that were preferred and thus highly trained on tension are activated with the onset of the MPT even in the presence of small compressive stresses. Hence, ε_a increases and ε_c decreases rapidly with cooling in the range of 40→24°C. As there is no compensation in energy for the activation of these variants, the transition range starts at 18°C and continues down to around 0°C, where ε_a tends toward ε_q (and similarly ε_c tends toward ε_p). In the temperature range 40→24°C $\varepsilon_p \approx \varepsilon_q$ (or $\varepsilon_a \approx \varepsilon_c$) and hence the strain distribution is essentially close to isotropic, similar to the case of the first run under zero load. In this range of temperatures, the behavior shifts from that dictated by tensile stresses to that dictated by compressive stresses. This shift can be described as the rotation of principal strain axes.

At the higher compressive stress of -1.3 ksi (Fig. 64(c)), the transition region shrank to a point at 18.5°C, i.e., its temperature range diminished to a single point, and shifted toward higher temperature. Below 18°C, the variants imposed by the compressive stress are dominant. At the still higher compressive stress of -2.6 ksi (Fig. 64(e)), the transition point moves higher on the temperature scale to 20°C, whereas at -5.9 ksi (not shown), it shifts out of the MPT range. From this applied stress on, the behavior of the strain hysteresis loop is essentially

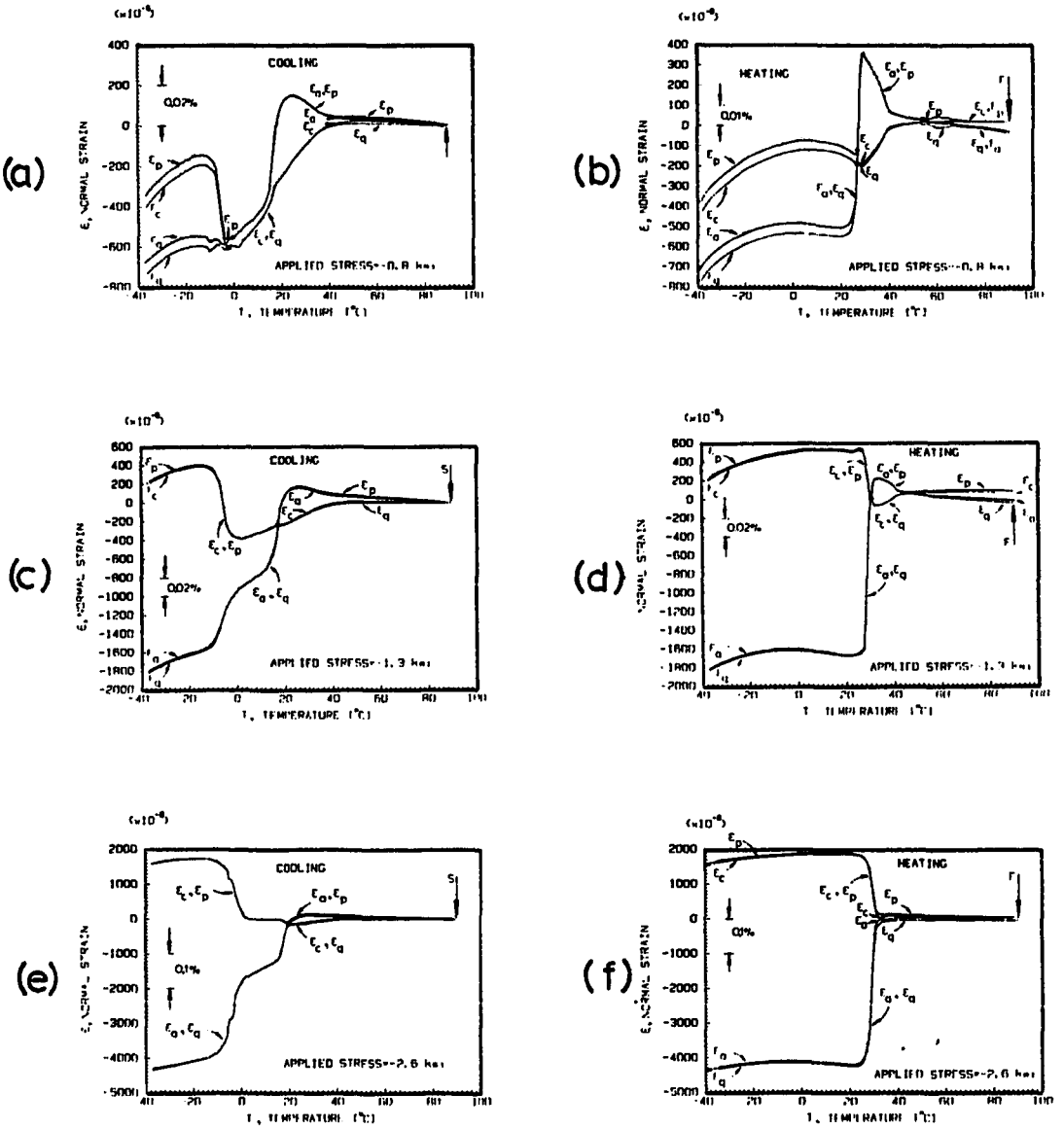


Figure 64. Normal strain versus temperature for sample tube P8-1 for (a) ε_a, ε_c, ε_p, and ε_q for applied compressive stress 0.8 ksi, (c) the same for applied compressive stress 1.3 ksi, and (e) the same for applied compression stress 2.6 ksi, all for cooling; and (b), (d), and (f) for heating, respectively. In many cases, the curves for the principal strains (ε_p and ε_q) coincide with the curves for axial and circumferential strains (ε_a and ε_c, respectively), as indicated

as shown in Figs. 62 (c) and (d), and the sample is trained for preferred activation of variants under an applied compressive stress.

It was found by Hsu [80] that Ni-Ti wires showed substages (wiggles) in their (axial) strain hysteresis loops on cooling. When wires made of the same material were tested in the same way after millions of transformation fatigue cycles (TFC), the strain hysteresis loops showed additional wiggles (Figs. B1, and B2 in ref. [80]). A heat treatment at 400°C for 1 hour did not change greatly the appearance of the hysteresis loops for the uncycled wires. On the other hand, the heat treatment eliminated some of the wiggles from the strain hysteresis loop on cooling in the case of TFC wires (Figs. B6 and B7 in ref. [80]). The appearance of wiggles in strain hysteresis loops of uncycled wires was attributed by Hsu to the B2→R transformation [80], and the additional wiggles in the hysteresis loops of TFC wires to the increase in dislocation density, which affects M_S (but not T_R) and hence the nature of the B2→R, B2→B19, and R→B19 crystallographic transformations [80, 87]. In Fig. 65, two sets of strain hysteresis loops are shown for: (a), (b) and (c), the circumferential strain behavior at 3.2, 1.0 and 0.5 ksi, respectively, on tension in the unloading pattern sequence, and (d), (e) and (f), the axial strain behavior at -5.9, -0.3, and 0^- ksi on compression in the unloading pattern sequence. The development of wiggles is clearly demonstrated. Thus, wiggles occur over a large portion of the hysteresis loop on cooling upon simple release of load. The main point here is that the

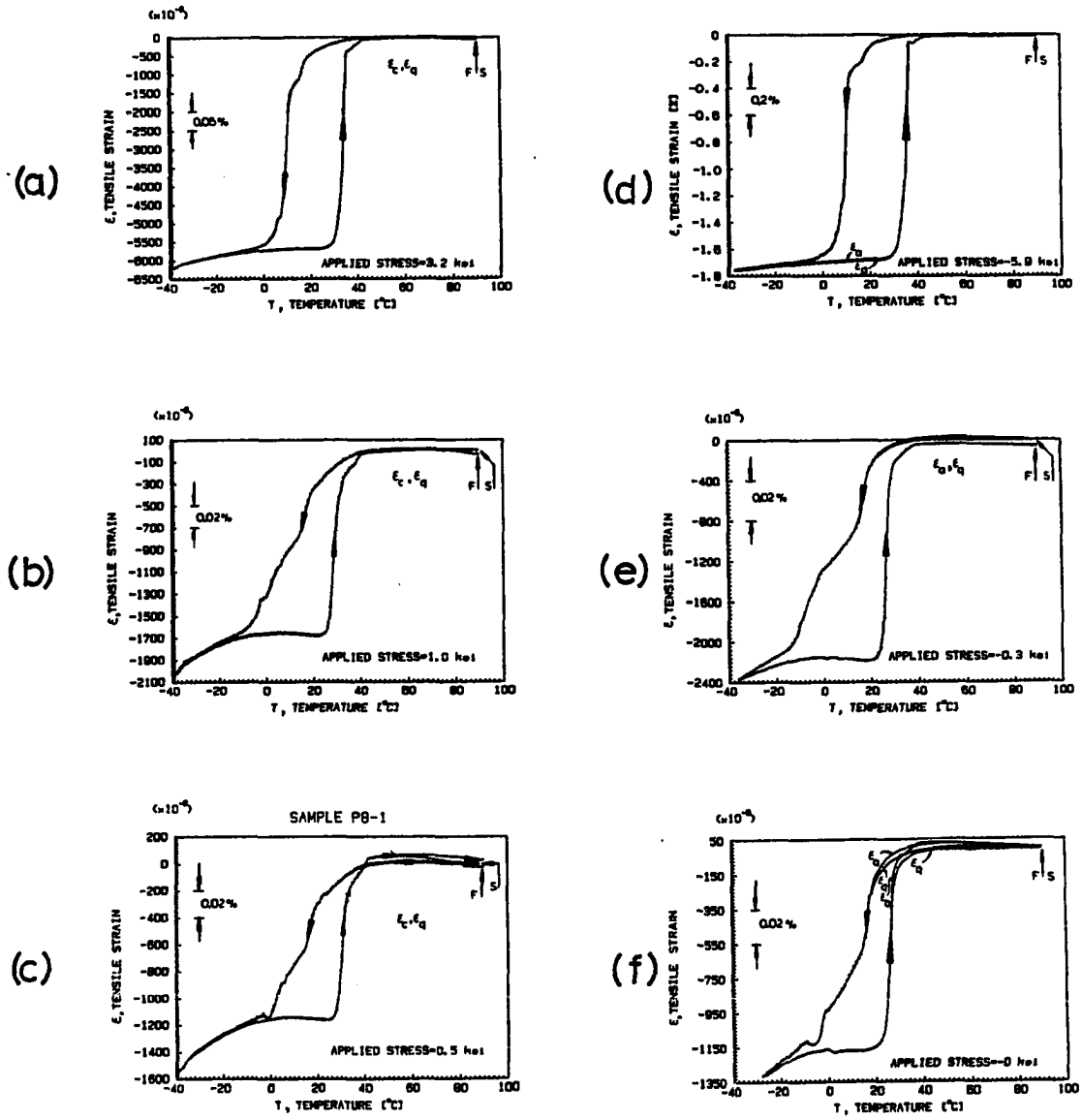


Figure 65. Strain-temperature curves. (a), (b) and (c): circumferential strain hysteresis loops for 2.3, 1.0 and 0.5 ksi applied tensile stress (in that order), respectively. (d), (e), and (f): axial strain hysteresis loops for -5.9, -0.3, and 0⁻ ksi compressive stress (in that order), respectively. In both cases, substages were developed on cooling which increased in magnitude with decreasing applied stress

strain hysteresis loop shape is defined by stress-training. The ability to erase the memory gained during forward transformation under applied stress may be affected by TFC.

3.4 Determination of the Transformation Volume Change

The MPT exhibits a small, but nevertheless non-zero volume change. This volume change can be measured by different methods, like lattice parameter measuresments [32, 34, 94], hydrostatic weighing [95] and strain measurements [95]. The last method was used to determine the volume change on transformation of the annealed square cross-sectional sample P7-1. The volume change on transformation was determined as a function of applied stress, including zero stress. To each side of P7-1, a strain gage was attached. Three of them measure in the three mutually perpendicular directions z, x, and y (sides 2, 3, and 4, respectively) the respective strains $\varepsilon_a^{(2)}$, $\varepsilon_t^{(3)}$ and $\varepsilon_t^{(4)}$, where z is along the sample axis. Side 1 contained a long narrow strain gage in the sample axis direction (measuring $\varepsilon_a^{(1)}$); thus, this strain gage integrated strain values over a larger area of the sample.

Fig. 66 shows the transformation strains indicated by the four strain gages. All the MPT strains are identical at zero applied stress, and are positive. Moreover, at zero load the shapes of the strain hysteresis loops (not shown) in the different directions are nearly identical, thus confirming that:

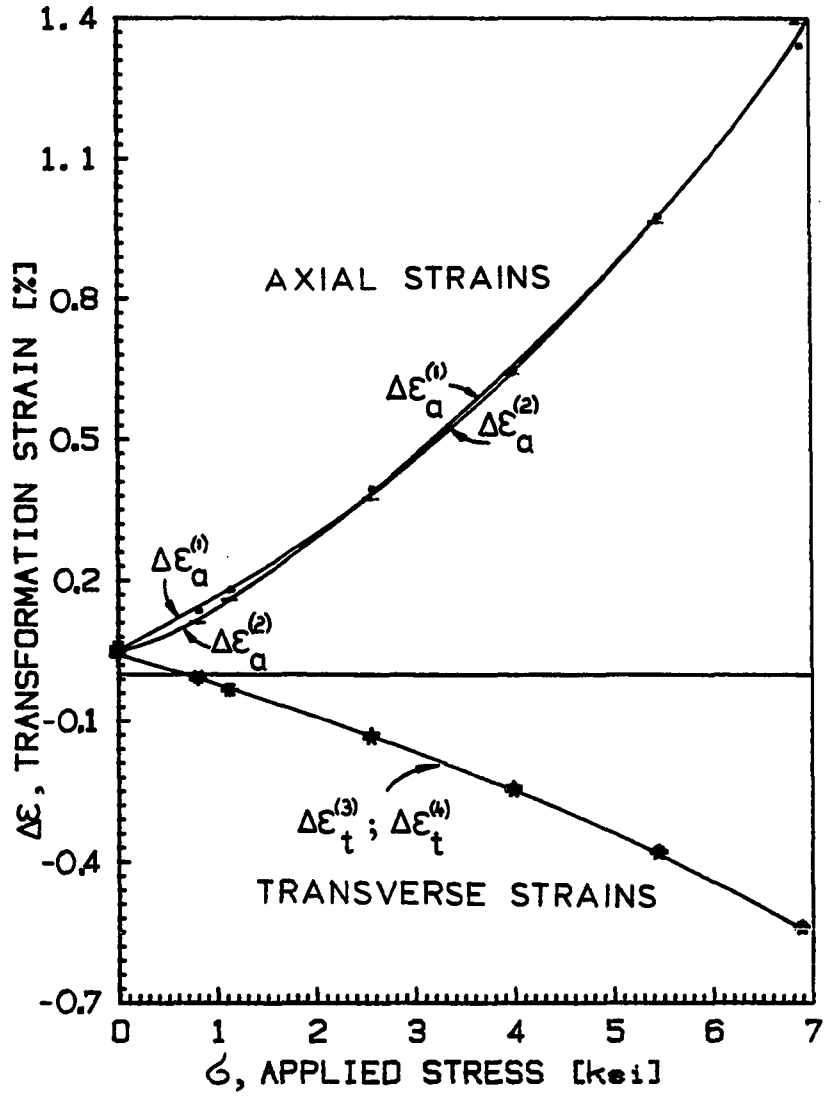


Figure 66. MPT strains as a function of applied tensile stress for the P7-1 square cross-sectional bar

(a) The transformation on the first run in absence of applied stress is isotropic.

(b) The relative volume change of virgin material at zero applied stress on cooling is positive.

Further, the MPT strains as a function of applied stress indicate that:

(c) There is an increase in the axial strain with increasing tensile stress (sample elongates axially) and a simultaneous increase in magnitude for the transverse strain with increasing tensile stress (sample contracts in the transverse plane), as is observed also for the tubular samples P8-1 (Fig. 54) and PO-4 (Fig. 60).

(d) The two transverse MPT strains are identical, and moreover they have the same general shape (not shown) over the entire transformation temperature range, both for forward and reverse transformations.

From the forward transformation strains, the fractional volume change is calculated using the relation:

$$\frac{\Delta V}{V_o} = \Delta\varepsilon_a^{(2)} + \Delta\varepsilon_t^{(3)} + \Delta\varepsilon_t^{(4)} \qquad (11)$$

where ΔV is the volume change on MPT and V_o is the HTP volume. All of the strain gages were zeroed at room temperature. Then, the sample was heated to 145°C at zero load, and for the loaded runs the load was applied. Then, the sample was cycled through the transformations on cooling and heating, and returned to 145°C, whereupon the new load was

applied. The loads were increased progressively from run to run. The effect of strain due to the applied load itself (similar to that shown for sample P8-1 in Fig. 48) was subtracted from each curve, such that each curve is at zero fractional volume change at 145°C. Thus, V_o is the volume of the HTP at 145°C, corrected to zero stress. As concerns thermal contraction upon cooling from 145°C to M_S, the strain gages are compensated for this such that the correction over this temperature interval is less than about 80×10^{-6} strain units.

The fractional volume change on cooling is shown in Fig. 67 as a function of applied stress. The zero stress value of fractional volume change on cooling is $\Delta V/V_o = 0.145\%$, in good agreement with those measured by Hsu [80] on cast sample by hydrostatic weighing (0.19%) and by strain measurements, based on the assumption that the fractional volume change can be approximated by $\Delta V/V_o = 3/2(\Delta\varepsilon_a + \Delta\varepsilon_c)$. It is, however, not in agreement with fractional volume change calculations from lattice parameter data from which the fractional volume change on cooling is negative [95].

The fractional volume $\Delta V/V_o$ on cooling (Fig. 68(a)), and on heating (Fig. 68(b)) was plotted versus sample temperature for the different applied tensile loads. For the cooling part of the figure, for zero applied stress (first run) the relative volume continuously increases with decreasing temperature from 50° down to -10°C. When stress is applied, however, there is at first a decrease in volume in the temperature range from 50 to 35°C, followed by an increase with further temperature drop.

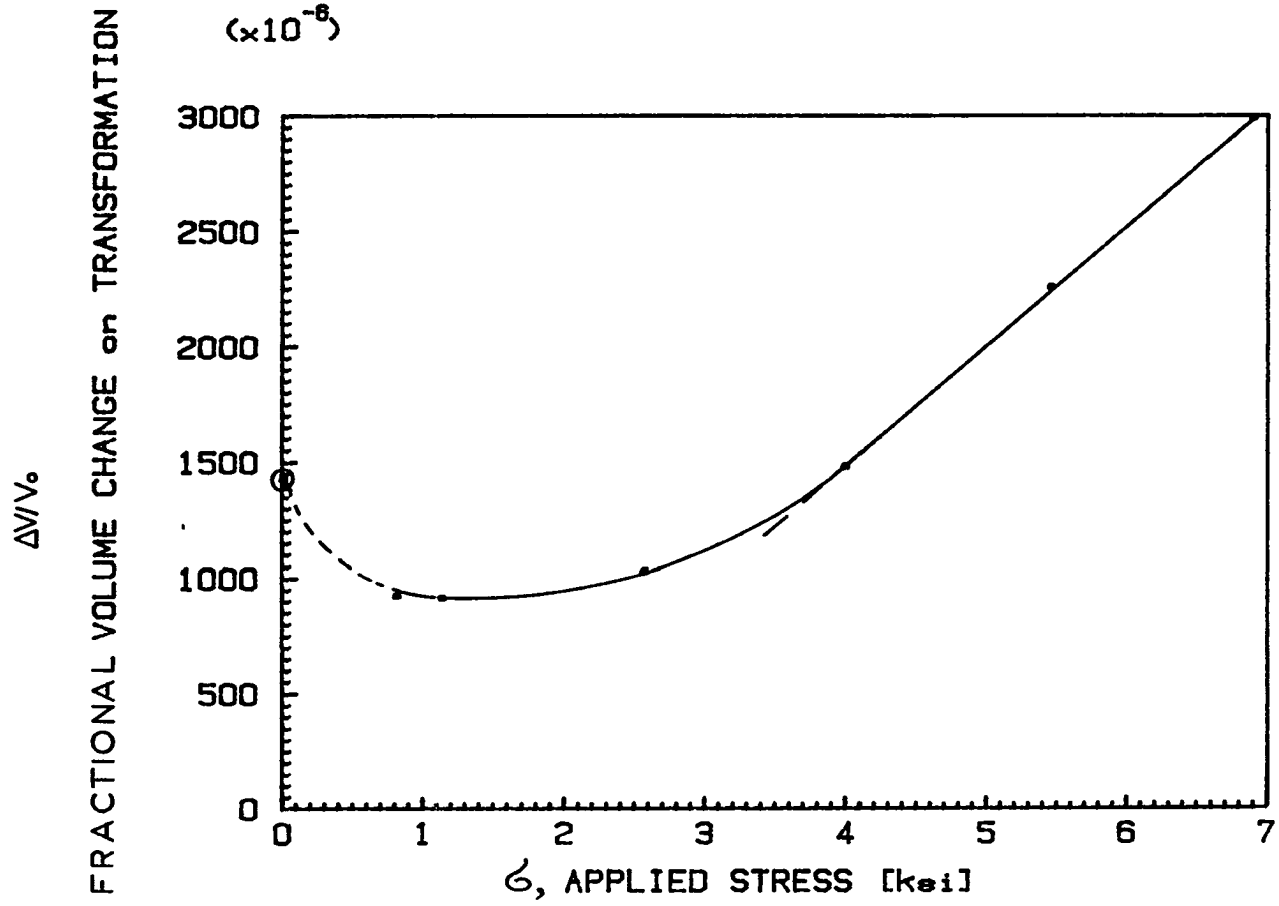


Figure 67. Fractional volume change on forward transformation as a function of applied stress for sample P7-1

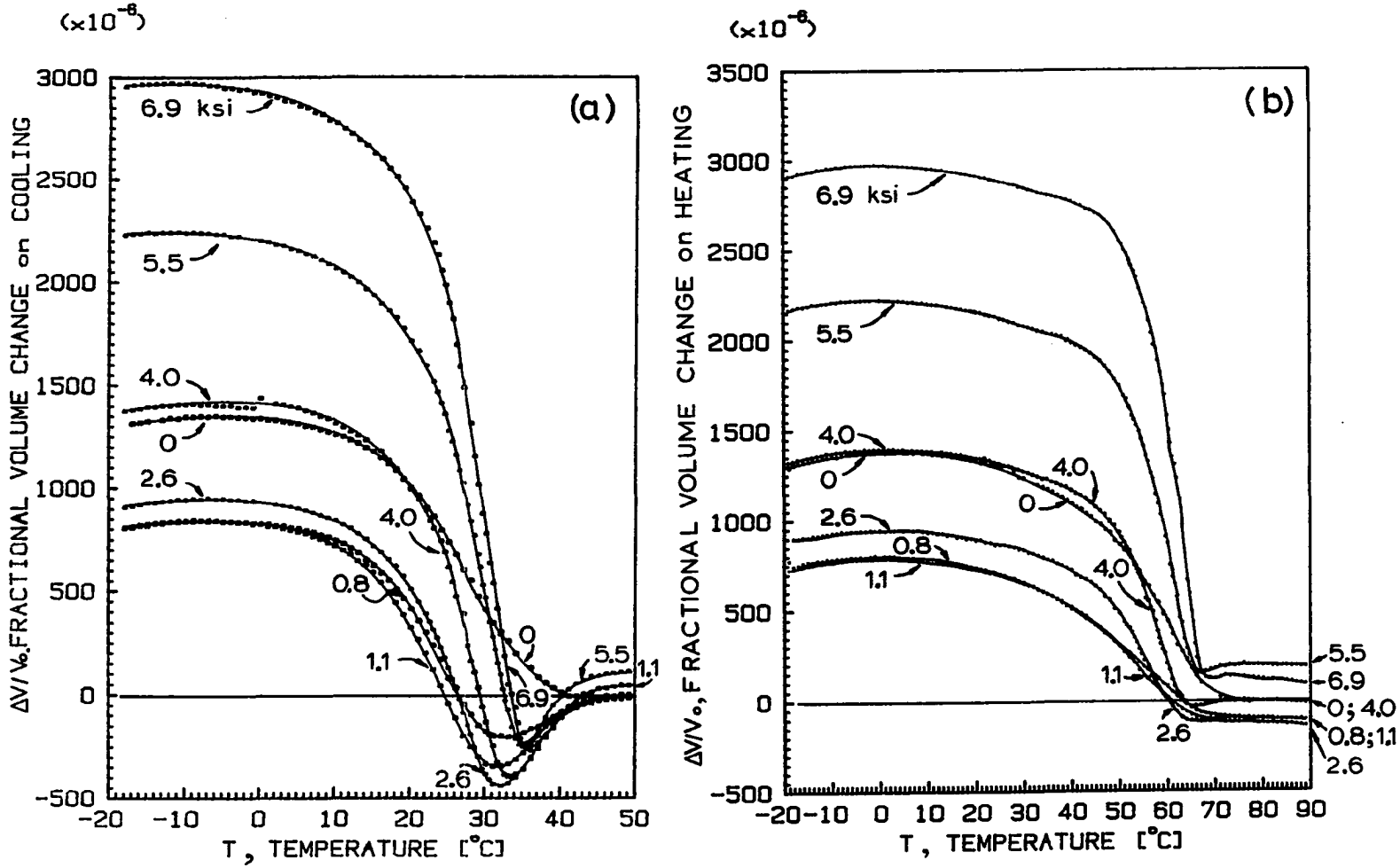


Figure 68. Fractional volume change during transformation as a function of sample temperature for the various applied tensile stresses, (a) on cooling, and (b) on heating

It was pointed out by Ling and Kaplow that the B2→R transformation (from the b.c.c. to a rhombohedral lattice) is accompanied by a negative dilation (contraction) and generally takes place at higher temperatures than the MPT [87]. Thus, the fractional volume decrease betwen 50 and 35°C for loaded runs (Fig. 68(a)) can be interpreted as the volume change induced by the B2→R transformation, i.e., a premartensitic effect and not related directly to the MPT as such. The curves of fractional volume change on heating do not show this anomaly (Fig. 68(b)). When the differential strain curves (section 3.1.2) were analyzed, it was found that the strain hysteresis loop is double activated on cooling under applied load (e.g., Fig. 29), but only single activated for the first run when no stress was applied (Fig. 27). If this double activation is caused by the B2→R transformation, then, as it was shown, it is clearly applied stress dependent. Thus, the different behavior of fractional volume change on cooling for the zero-load run in Fig. 68(a) is consistent with the other observations.

The total fractional volume change on cooling as a function of applied stress (Fig. 67) exhibits several regimes of different behavior. While for applied stresses below 4 ksi the fractional volume change dependence on applied stress is small, above 4 ksi it shows a linear increase with applied stress. Such a behavior is not expected on the basis of the PAM. Based on the PAM, the fractional volume change of transformation should be independent of the applied stress, as it is unaffected

by the orientation of the distortion ellipsoids. A possible explanation of this anomaly which still leaves the PAM intact is as follows.

The applied uniaxial tensile stress is constant and equal at the HTP and LTP (neglecting changes in cross-section). In the three dimensional axis system, with the axes being σ_a, ε, and T the total process of loading at HTP, cooling to LTP, and measuring the strain can be visualized as a three step process:

(a) Upon loading at the HTP, the material behaves as a perfect elastic body, and elongates in the axial (tensile stress) direction along a line given in the σ-ε plane (for $T>M_S$) by the HTP Young's modulus. This small elastic elongation was cancelled out from the data points of Fig. 67.

(b) Lowering the temperature (on cooling) the MPT takes place and contributes the transformation strain to the total observed strain.

(c) At the LTP, the stress-strain behavior (σ-ε plane at $T\approx M_F$) is dictated by a line given by the Young's modulus of the LTP, and the strain readjusts itself such that the applied stress remains the same as was at the HTP. (This readjustment is small, depending only on the difference between the two Young's moduli, at the HTP and LTP, respectively.)

The above sequence is reasonable provided the applied stress, σ_a, is lower than the elastic limit. From [96], however, $\sigma_y(\text{HTP})>>\sigma_y(\text{LTP})$.

Now, we assume that σ_y(LTP) is approximately 4 ksi. Then, up to 4 ksi, $\Delta\varepsilon_a$ increases with σ_a (and $\Delta\varepsilon_t$ <0 and increases in magnitude), hence

$$\frac{\Delta V}{V_o} = \Delta\varepsilon_a + 2\Delta\varepsilon_t = \Delta\varepsilon_a - 2|\Delta\varepsilon_t| \approx \text{const}$$

according to the schematic description above ((a), (b), and (c)). However, above 4 ksi, (c) doesn't hold. The strain will readjust itself only up to 4 ksi along the straight line of the LTP Young's modulus, from which on it will flow without increase in stress, followed by an increase in both strain and stress along a straight line which has a smaller slope than that of the LTP Young's modulus line (See Fig. 91 in ref. [29]). Hence, above 4 ksi, the increase in strain in the axial direction will be much larger than below 4 ksi with increasing applied stress, and the major contribution to it will be that of the "plastic" strain at LTP for σ_a above σ_y(LTP). This increase in strain is approximately linear in σ_a. (This "plastic" strain recovers upon heating to the HTP due to the SME.) The $\Delta\varepsilon_t$'s, however, are not affected by plastic strain contribution as the applied stress is uniaxial and has zero value in the transverse directions. Hence, they continue to be equal to the transformation strain in the transverse direction.

As still:

$$\frac{\Delta V}{V_o} = \Delta\varepsilon_a - 2|\Delta\varepsilon_t|$$

and as:

$$\Delta\varepsilon_a = \Delta\varepsilon_a^{tr} + \Delta\varepsilon_a^{p}$$
$$\Delta\varepsilon_t = \Delta\varepsilon_t^{tr} \approx -\frac{1}{2}\Delta\varepsilon_a^{tr}$$

where $\Delta\varepsilon_a$ and $\Delta\varepsilon_t$ are the total strains in the axial and transverse directions, respectively. $\Delta\varepsilon_a{}^{tr}$ and $\Delta\varepsilon_t{}^{tr}$ are the transformation strains in the above directions, and $\Delta\varepsilon_a{}^{p}$ is the axial plastic strain. It follows that for $\sigma_a > \sigma_y$(LTP)

$$\frac{\Delta V}{V_o} \approx \Delta\varepsilon_a^{p}$$

as is shown in Fig. 67, for $\sigma_a > 4$ksi.

4. CONCLUSIONS

The transformation strain dependence on uniaxially applied tensile and compressive stress in equiatomic Ni-Ti samples was investigated. The experimental results were discussed in terms of a preferred activation model (PAM) and the phenomenological WLR theory. From these studies, the following conclusions are drawn:

(a) In an annealed sample on first run and under zero load condition, the various crystallographically equivalent variants of the martensite phase transformation (MPT) are randomly activated. The strain-temperature hysteresis loops are similar in appearance for strains in all directions relative to the sample axis.

(b) Applied stress on thermal cycling preferentially activates certain of the martensite variants. The variants that are able to contribute the maximum release of stress in the applied stress direction tend to be preferred. Thus, the preferentially activated variants are those for which the conjugate total distortion ellipsoid induced by the MPT will be aligned along the applied stress direction. For applied tensile stress, the major axis of the ellipsoid will be in the uniaxial stress direction, while for compressive stress the minor axis will lie in the axial direction. Thus, the axial transformation strain on cooling increases with applied tensile stress and is positive.

Also, the transverse strains increase in magnitude with increasing tensile stress levels and are negative. Moreover, the major and minor principal strains are aligned with the axial and transverse directions in the sample, respectively, and increase in magnitude with increasing applied stress.

(c) For applied uniaxial compressive stress, the results are reversed as compared to the case of uniaxial tension (case (b) above). Maximum release in stress occurs if the distortion ellipsoids of the preferentially activated martensite variants are aligned with their minor axis (corresponding to contraction) along the applied stress direction. Thus, it was found that the axial transformation strain on cooling increases in magnitude with applied compressive stress and it is negative, i.e., the sample contracts in its axis. The transverse strains increase with increasing compressive stress and are positive, i.e., the sample expands in its transverse directions. Thus, the major and minor principal strains are aligned with the transverse and axial directions in the sample, respectively, and increase in magnitude with increasing applied uniaxial compressive stress.

(d) The axial transformation strain on cooling at ±7.5 ksi applied stress was found to be ~2.5%. Calculation of the numerical values of the semi-axes of the total distortion ellipsoid shows that the upper theoretical limit of the axial transformation strain is 5.4%. The transformation strain approaches this limit

when the major axes of all the distortion ellipsoids are aligned in the direction of the applied uniaxial tensile stress, i.e., when only the most preferred variant operates. The above numerical value was calculated by applying the WLR theory to the Ni-Ti case and using the lattice parameters reported by Otsuka et al. [33].

(e) The PAM predicts that a time sequence relation exists in the activation of variants under applied load. The variants that are most preferred will be activated first, simultaneously with the onset of the MPT. This sequential activation is the major contributor to the training of the sample. The strain during the not loaded run, following a high applied stress run, will, to some extent, behave at the first stages of the MPT as if the run was loaded. Hence, the effect of loading sequence on the transformation strain is very important.

(f) The effect of prior thermomechanical history was studied, and interpreted in terms of the extent to which the activation of martensite variants was random or preferred. The detailed nature of the thermomechanical history was observed to exert a strong influence on the shapes of the strain-temperature hysteresis loops, particularly at lower applied stresses. At high stresses, the formation of the particular martensite variants giving rise to the appropriate strains along the stress axis is especially energetically favorable. As a result, the

ε-T curves on cooling and heating are monotonic and well-separated. But at low stresses, other factors compete successfully and cause the hysteresis loops to exhibit gross departures from the simple behavior. This was observed to occur, for example, when successive runs were first conducted at higher and higher tensile stresses and then at lower and lower tensile stresses, going through zero stress, and then following the same pattern for compression. The ε-T hysteresis loops, thus obtained for stresses near zero, showed a complex behavior, which is described as a twisted hysteresis loop for which the cooling and heating curves crossed one another. The additional factors that cause this complexity at lower applied stresses are thought to be related to the need for various variants to be self-accommodating. When self-accommodation is achieved, the internal stresses (resulting for example, from volume and shear constraints) from various variants tend to cancel one another.

(g) From strain gage measurement data, the fractional volume change of phase transformation on cooling was found to be positive and to have the value of 0.15%.

5. REFERENCES

1. C. M. Wayman, Martensitic Transformations, in Science and Technology of Zirconia, Advances in Ceramics, Vol. 3, edited by A. H. Heuer and L. W. Hobbs (The American Ceramic Soc., Inc., Columbus, Ohio, 1981).

2. E. C. Bain, Trans. AIME, 70, 25 (1924).

3. M. S. Wechsler, D. S. Lieberman, and T. A. Read, Trans. AIME 197, 1503 (1953)

4. G. V. Kurdyumov and A. G. Khandros, Zh. Tekh. Fiz., 19 (1949).

5. C. M. Wayman and K. Shimizu, Metal Science J., 6, 175 (1972).

6. C. M. Wayman, J. Metals, p. 129, June 1980.

7. R. Banks, Nitinol Heat Engines, in Shape Memory Effects in Alloys, edited by J. Perkins (Plenum: New York, 1975).

8. Advertising Brochure, ME-004, Raychem Corporation-Menlo Park, CA, 1971.

9. G. F. Andreasen and R. E. Morrow, Am. J. Orthodontics, 73(2), 142 (1978).

10. H. Wagner and C. Jackson, Mater. Eng., 70(7), (1969).

11. L. C. Chang and T. A. Read, Trans. AIME, 189, 47 (1951).

12. K. Otsuka and K. Shimizu, Scripta Met., 4, 469 (1970).

13. C. M. Wayman, Scripta Metallurgica, 5, 489 (1971).

14. P. Wollants, M. DeBonte, L. Delaey, and J. R. Roos, Z. Metallkde, 70, 298 (1979).

15. K. Enami, S. Nenno, and Y. Minato, Scripta Metallurgica, 5, 663 (1971).

16. F. E. Wang, W. J. Buehler, and S. J. Pickart, J. Applied Physics, 36, 3232 (1965).

17. Z. S. Basinski and J. W. Christian, Acta Met., 2, 161 (1954).

18. H. Margolin, E. Ence, and J. P. Nielsen, Trans. AIME, 197, 243 (1953).

19. D. M. Poole and W. Hume-Rothery, J. Inst. Metals, 83, 473 (1954).

20. G. R. Purdy and J. G. Parr, Trans. AIME, 221, 636 (1961).

21. K. Hirano and K. Ouchi, J. Japan Inst. Metals, 32, 613 (1968).

22. T. Suzuki, Trans. Japan Inst. Metals, 14, 31 (1973).

23. W. J. Buehler and R. C. Wiley, Trans. ASM, 55, 269 (1962).

24. R. Scholl, D. J. Larson, Jr., and E. J. Freise, J. Applied Physics, 39, 2186 (1968).

25. A. G. Rozner, E. F. Heintzelman, W. J. Buehler, and J. V. Gilfrich, American Society of Metals, Trans. Quart. 58, 415 (1965).

26. M. S. Wechsler, H. H. Baker, A. J. Bevolo, E. D. Gibson, and F. C. Laabs, Microstructure and Microchemistry of Equiatomic Nickel-Titanium, Report IS-4797. Ames Laboratory USDOE, Iowa State University, Ames, Iowa (1982).

27. R. J. Wasilewski, S. R. Butler, and J. E. Hanlon, Metals Science J., 1, 104 (1967).

28. M. H. Mueller and H. W. Knott, Trans. AIME, 227, 674 (1963).

29. C. M. Jackson, H. J. Wagner, and R. J. Wasilewski, 55-Nitinol-The Alloy with Memory: Its Physical Metallurgy, Properties, and Applications, Report NASA-SP5110. (National Aeronautics and Space Administration, Washington, D.C., 1972).

30. D. P. Dautovich and G. R. Purdy, Can. Metal. Quart., 4, 129 (1965).

31. M. J. Marcinkowski, A. S. Sastri, and D. Koskimaki, Phil. Mag., 18, 945 (1968).

32. T. V. Philip and P. A. Beck, Trans. AIME, 209, 1269 (1957).

33. K. Otsuka, T. Sawamura, and K. Shimizu, Physica Status Solidi (a), 5, 457 (1971).

34. G. M. Michal and R. Sinclair, Acta Crystallogr., B37, 1803 (1981).

35. M. A. Jawson and J. A. Wheeler, Acta Crystallogr., 1, 216 (1948).

36. H. A. Mohamed and J. Washburn, J. Materials Science, 12, 469 (1977).

37. R. J. Wasilewski, Met. Trans. A., 6A, 1405 (1975).

38. F. Foerster and E. Scheil, Z. Metallk., 32, 165 (1940).

39. R. F. Bunshah and R. F. Mehl, Trans. AIME, 197, 1251 (1953).

40. O. Kirsement, E. Houdremont, and F. Wever, Rev. Met., 51, 401 (1954).

41. D. Goldstein, A Source Manual for Information on Nitinol and NiTi, NSWC TR 80-59 (Naval Surface Weapons Center, Silver Springs, Maryland, 1980).

42. R. J. Wasilewski, Scripta Met., 5, 131 (1971).

43. G. D. Sandrock, A. J. Perkins, and R. F. Hehemann, Met. Trans., 2, 2769 (1971).

44. K. Otsuka, T. Sawamura, K. Shimizu, and C. M. Wayman, Met. Trans., 2, 2583 (1971).

45. C. M. Wayman, I. Cornelis, and K. Shimizu, Scripta Met., 6, 115 (1972).

46. K. H. Eckelmeyer, Scripta Met., 10, 667 (1976).

47. W. Schwenk and J. Huber, Sampe Quarterly, 5, 17 (1974).

48. R. J. Wasilewski, S. R. Butler, J. E. Hanlon, and D. Worden, Met. Trans., 2, 229 (1971).

49. J. V. Gilfrich, Adv. X-ray Anal., 6, 74 (963).

50. W. J. Buehler and F. E. Wang, Ocean Engng., 1, 105 (1968).

51. J. E. Hanlon, S. R. Butler, and R. J. Wasilewski, Trans. TMS-AIME, 239, 1323 (1967).

52. F. E. Wang, Proc. of the First Intern. Conf. on Fracture, Vol. 2, Sendai, Japan, 1965.

53. J. Perkins, Met. Trans., 4, 2709 (1973).

54. J. Perkins, Scripta Met., 9, 121 (1975).

55. L. Kaufman and M. Cohen, Prog. Met. Phys., 7, 165 (1958).

56. R. J. Wasilewski, Scripta Met., 5, 207 (1971).

57. G. K. Bansal and A. H. Heuer, Acta Met., 20, 1281 (1972).

58. N. Nakanishi, T. Mori, S. Miura, Y. Murakami, and S. Kachi, Phil. Mag., 28, 277 (1973).

59. H. C. Tong and C. M. Wayman, Met. Trans. A., 6A, 29 (1975).

60. C. M. Wayman and H. C. Tong, Scripta Met., 11, 341 (1977).

61. I. Cornelis and C. M. Wayman, Scripta Met., 10, 359 (1976).

62. H. C. Tong and C. M. Wayman, Scripta Met., 8, 93 (1974).

63. L. Delaey, R. V. Krishnan, H. Tas, and H. Varlimont, J. Materials Science, 9, 1521 (1974).

64. W. J. Buehler, J. V. Gilfrich, and R. C. Wiley, J. Applied Physics, 34, 1475 (1963).

65. R. R. Hasiguti and K. Iwasaki, J. Applied Physics, 39, 2182 (1968).

66. S. P. Gupta, Phys. Stat. Sol., a13, K5 (1972).

67. R. J. Wasilewski, Scripta Met., 9, 417 (1975).

68. H. Pops, Met. Trans., 1, 251 (1970).

69. H. Pops and L. Delaey, Trans. AIME, 242, 1849 (1968).

70. J. Perkins, Metallography, 7, 345 (1974).

71. W. Arneodo and M. Ahlers, Scripta Met., 7, 1298 (1973).

72. M. W. Burkart and T. A. Read, Trans. AIME, 197, 1516 (1953).

73. Z. S. Basinski and J. W. Christian, Acta Met., 2, 101 (1954).

74. J. Perkins, Scripta Met., 8, 1469 (1974).

75. V. N. Khachin and L. A. Solovev, Phys. Stat. Sol.(a), 30, 671 (1975).

76. H. Tas, L. Delaey, and A. Deruyttere, J. Less Common Metals, 28, 141 (1972).

77. A. Nagasawa, K. Enami, Y, Ishimo, Y. Abe, and S. Nenno, Scripta Met., 8, 1055 (1974).

78. G. F. Bolling and R. H. Richman, Scripta Met., 4, 539 (1970).

79. C. H. Hsu and M. S. Wechsler, in Proc. Int. Conf. on Solid-Solid Phase Transformations, American Society for Metals, Carnegie-Mellon University, Pittsburgh, August 10-14, 1981.

80. C. H. Hsu, Ph. D. Thesis, Iowa State University, 1982.

81. G. R. Edwards, J. Perkins, and J. M. Johnson, Scripta Met., 9, 1167 (1975).

82. M. T. Podob, W. A. Johnson, and S. H. Reichman, Hot Isostatically Pressed Powder Metallurgy Nitinol Wire, in Proc. of Nitinol Heat Engine Conference, Silver Spring, MD, 26-27 Sept., 1978.

83. Metals Handbook, 9th Edition, Vol. 1, Properties and Selection: Irons and Steels (American Society for Metals, Metals Park, Ohio 1978).

84. J. W. Dally and W. F. Riley, Experimental Stress Analysis, 2nd Edition, McGraw-Hill Book Comp., N.Y., 1978.

85. Manual for V/E-20A Digital Strain Indicator, Vishay Measurements Group, Raleigh, NC., (1978).

86. R. M. Banks and Associates Machine Development, 1805 Eastshore Highway, Berkeley, CA 94710, private communication (1981).

87. H. C. Ling and R. Kaplow, Mat. Sci. Eng., 51, 193 (1981).

88. M. S. Wechsler, Iowa State University, private communication (1981).

89. H. A. Mohamed and J. Washburn, Met. Trans., 7A, 1042 (1976).

90. C. M. Wayman, Deformation, Mechanism and Other Characteristics of Shape Memory Alloys, in Shape Memory Effects in Alloys, edited by J. Perkins, (Plenum: New York, 1975).

91. F. J. Wyant, M.S. Thesis, Iowa State University, 1980.

92. H. A. Mohamed, J. Materials Science, 13, 2729 (1978).

93. A. G. Rozner and R. J. Wasilewski, J. Inst. Metals, 94, 169 (1966).

94. R. F. Hehemann and G. D. Sandrock, Scripta Metall., 5, 801 (1971).

95. C. H. Hsu, M. S. Wechsler, and H. Diehl, Transformation Volume Change in Equiatomic Nickel-Titanium by Hydrostatic Weighing, Report IS-4799, Ames Laboratory USDOE, Iowa State University, Ames, Iowa (1982).

96. H. A. Mohamed, Ph. D. Thesis, University of California, Berkeley, CA (1976).

97. D. Bradley, J. the Acoustical Soc. America, 37, 700 (1965).

98. D. Turnball and J. C. Fisher, J. Chem. Phys., 17, 71 (1949).

99. D. S. Lieberman, M. S. Wechsler, and T. A. Read, J. Applied Physics, 26 (4), 473 (1955).

100. K. M. Knowles and D. A. Smith, Acta Met. 29, 101 (1981).

101. Z. Nishiyama, Martensitic Transformations, Academic Press, 1978.

102. J. W. Christian, The Theory of Transformations in Metals and Alloys, Pergamon Press, 1965.

103. E. O. Hall, Twinning and Diffusionless Transformations in Metals, Butterworth Sci. Pub., London 1954.

104. K. M. Knowles, Phil. Mag. 45 (3), 357 (1982).

105. M. A. Jawson and D. B. Dove, Acta Crystallogr., 13, 232 (1960).

106. R. W. Cahn, Acta Met. 1, 49 (1953).

107. S. P. Gupta and A. A. Johnson, Trans. Japan. Inst. Metals, 14, 292 (1973).

6. ACKNOWLEDGEMENTS

I wish to express my deep gratitude to Professor Monroe S. Wechsler, not only for his instruction and guidance, which were always at the highest professional level and without which this work never could be written, but also for his human heartedness to any problem I may have encountered during the years we spent together at Ames Laboratory.

I would like to acknowledge the contribution of the Advisory Committee to reading and correcting the manuscript and also for the moral support given by Professors R. A. Danofsky, F. X. Kayser, and D. M. Roberts.

I appreciate the kindness of the Special Metals Company of New Hartford, New York, and particularly W. A. Johnson, S. Reichman, and J. Pridgeon, who provided the powder metallurgical nickel-titanium starting stock.

Thanks are due to the Nuclear Research Centre-Negev, Israel Atomic Energy Commission for giving me the opportunity to leave for three years, and to Ames Laboratory, USDOE, for accepting me as a research assistant.

My acknowledgements to Dr. D. M. Martin and Mr. M. D. Rasmussen for their help in computer programming, to Mr. R. Z. Bachman for chemical analysis, to Mr. H. H. Baker for taking and preparing the micrographs, to Mr. T. Naig for his technical support, to Mr. D. W. Sailsbury for graphical aid, and to Mrs. L. P. Kilmer for the neat typewriting.

Last, but not least, I wish to express my heartfelt thanks to my family, my wife Tova, my children Ady and Lior, my sister Suzanne and my brother-in-law Benoit, for their outstanding support and understanding.

7. APPENDIX A: DIFFERENTIAL STRAIN HYSTERESIS LOOPS

7.1 The Mathematical Treatment and Some of the Physical Consequences

This mathematical formalism is essentially an outcome of the experimental observations. It is a method by which the macroscopic strain induced by the transformation can be characterized. To begin with, consider the axial strain hysteresis loop of a sample under zero external stress, as shown in Fig A1(a).

The loop consists of three major parts, where the subdivision criterion is the functional behavior of strain, ε_a, versus temperature, T; i.e., $\varepsilon_a = \varepsilon_a(T)$:

(a) Linear behavior, governed by the linear thermal coefficient of expansion of the sample material, α_{LTP}, for the low temperature phase. The strain in this region is given by:

$$\Delta\varepsilon_{LTP}(T) = \alpha_{LTP} \cdot \Delta T \qquad \text{(A-1)}$$

(b) Linear behavior governed by the linear thermal coefficient of expansion of the sample material at the high temperature phase configuration, α_{HTP}. The strain in this region is given by:

$$\Delta\varepsilon_{HTP}(T) = \alpha_{HTP} \cdot \Delta T \qquad \text{(A-2)}$$

Since the phases are different in these two extreme regions of the thermal cycle, $\alpha_{LTP} \neq \alpha_{HTP}$, and the two straight lines given by equations (A-1) and (A-2) have different slopes.

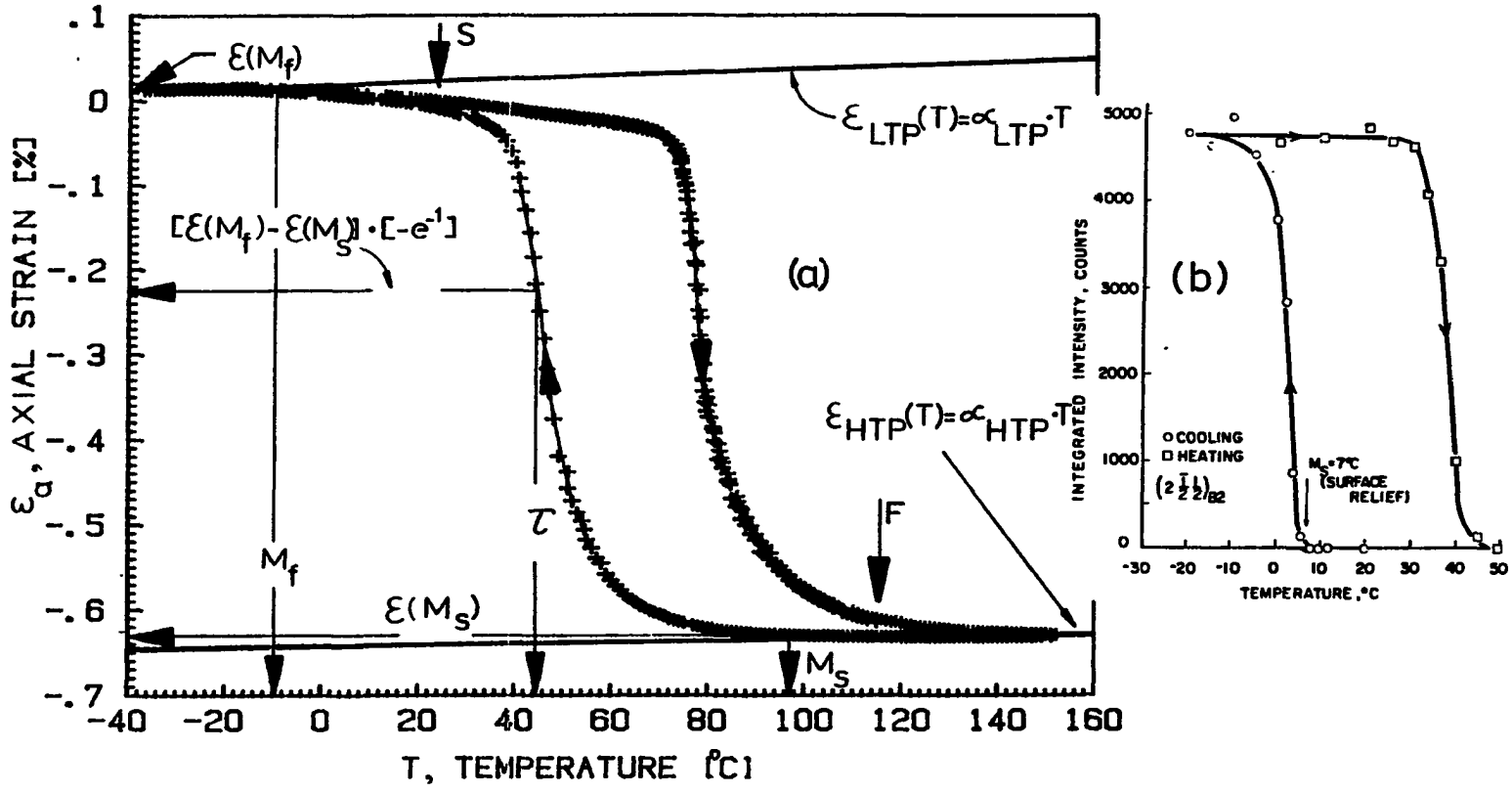


Figure A1. Phase transformation hysteresis loops. (a) Axial strain hysteresis loop of a sample under zero applied stress. The straight lines $\varepsilon_{LTP}(T)$ and $\varepsilon_{HTP}(T)$ are the thermal expansion induced strains below martensite finish, M_F, and above martensite start, M_S, temperatures, respectively. The figure shows the points $(M_S, \varepsilon(M_S))$ and $(M_F, \varepsilon(M_F))$, as well as the temperature constant τ and its corresponding strain value explicitly. (b) Variation in intensity of $1/2\{110\}_{B2}$ reflection with temperature. Zirconium-filtered molybdenum radiation, 36 kV-16 ma [43]

(c) The phase transformation region bounded by the above characterized straight lines. This portion of the loop, where the phase transformation induced gap between the two straight lines of thermal expansion is filled up is the subject of the following discussion.

Concentrating on the cooling part of the thermal cycle, we identify the martensite start, M_S, and the martensite finish, M_F, temperatures as the temperatures at which the strains deviate from those given in equations (A-2) and (A-1), respectively, as illustrated in Fig. A1(a). Later, another definition will be given for M_S and M_F. For the hysteresis loop of Fig. A1(a):

$$M_S = 370°K, \qquad \varepsilon(M_S) = -\ 6.3 \times 10^{-3}$$

$$M_F = 264°K, \qquad \varepsilon(M_F) = 0.1 \times 10^{-3}$$

To simplify the mathematical procedure and to cancel out the effect of zeroing the strain gage at a definite temperature, let's define the normalized transformation strain, f(T), as:

$$f(T) = \frac{\varepsilon(T) - \varepsilon(M_S)}{\varepsilon(M_F) - \varepsilon(M_S)}, \qquad M_F < T < M_S \tag{A-3}$$

Clearly,

$$f(T) = \begin{cases} 0 \text{ if } T = M_S \\ 1 \text{ if } T = M_F \end{cases} \tag{A-4}$$

Sandrock et al. investigated the martensitic phase transformation in Ni-Ti (~50 at. % Ni) by single crystal x-ray diffraction [43]. They

plotted (Fig. 11 in ref. [43]) the variation in intensity of the anomalous $1/2\{110\}_{B_2}$ reflection with sample temperature for the forward and reverse transformations (Fig. A1(b)). They state that for bulk samples this reflection is characteristic of the formation of martensite. Hence, assuming that the phase transformation is complete, i.e., in the LTP the sample consists of 100% martensite and no retained austenite, in Fig. A1(b) the ordinate can be expressed as percentage of martensite present instead of counts. From the similar behavior of curves A1(a) and A1(b), the same must be true of the ordinate of Fig. A1(a). Hence, one can look on f(T) in equation (A-3) not only as the normalized axial strain, but also as the relative amount of martensite present as a function of temperature. It should be noted, that this is not true, for example, of the sound velocity or the internal friction, Q^{-1}. These loops are not monotonic with temperature in the region of the forward or reverse transformations [65, 97]. As a consequence, both the $1/2\{110\}_{B_2}$ reflection intensity and the axial strain (and hence f(T)) are kinetic variables of the MPT. This will be discussed further below.

The immediate task now is to find a functional dependence of f(T) on material constants like M_S and M_F. Considering the sigmoidal shape of the cooling part of the hysteresis loop in Fig. A1(a), let us assume that f(T) can be given by the functional relation:

$$f(T) = g(y) \qquad \text{(A-5)}$$

where $y = \exp(-\Gamma)$, and Γ must satisfy the asymptotic conditions:

$$\Gamma = \begin{cases} 0 \text{ if } T \to T_S^- \\ +\infty \text{ if } T \to T_F^+ \end{cases} \qquad \text{(A-6)}$$

and where T_S^- and T_F^+ are the left handed and the right handed limits of the transformation regime, respectively. The usual way to designate these limiting start and finish transformation temperatures is by M_S and M_F, respectively. However, we will save the M_S and M_F notations for temperature values derived from the experimental strain curves, and thus leave the freedom to adjust the limiting points, T_S and T_F such that the analytical f(T) curve will give the best fit to the experimental data values.

If Γ is given by

$$\Gamma = \frac{T_S - T}{T - T_F} \qquad \text{(A-7)}$$

then this Γ satisfies the asymptotic conditions of equation (A-6), and moreover $\Gamma > 0$ provided $T_F^+ < T < T_S^-$.

Further, we define an intermediate function h(T) as

$$h(T) = 1 - \exp(-\Gamma) \qquad \text{(A-8)}$$

where Γ defined by equation (A-7).

In Fig. A2, the function h(T) is plotted versus the temperature for constant T_F(=260°K) and increasing $\Delta T = T_S - T_F$ (60, 80, 100, 120°K). From equation (A-8), a temperature constant τ can be defined as the temperature at which h(T) is within e^{-1} of its final value, i.e., where $\Gamma = 1$. But, from equation (A-7)

$$\Gamma = \frac{\tau - T_S}{T_F - \tau} = 1 \text{ yields } \quad \tau = \frac{T_S + T_F}{2} \qquad \text{(A-9)}$$

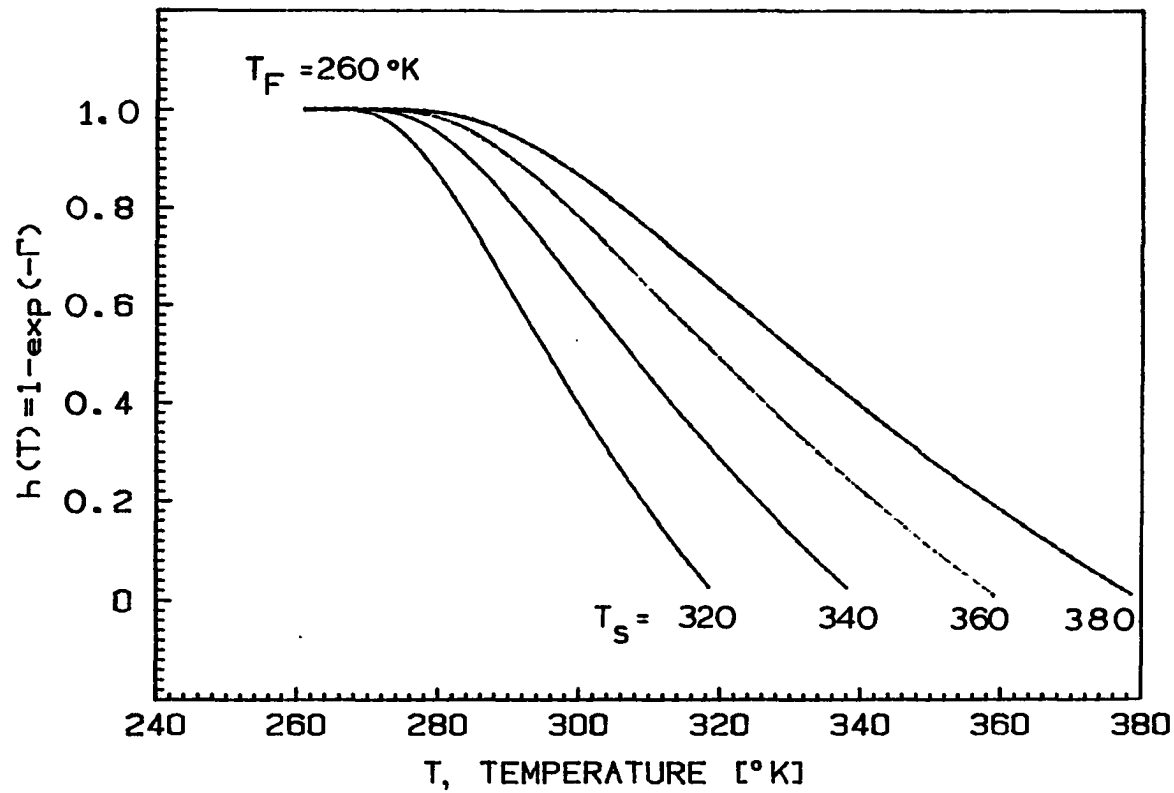


Figure A2. Family of curves of the intermediate function, $h(T)=1-\exp(-\Gamma)$, for T_F=260°K. and T_S=320, 340, 360 and 380°K

In the following, we assume that the "rate" at which the transformation progresses (i.e., the change of the relative amount of martensite present per unit temperature change) is proportional to -dh/dT. On the other hand, the rate of the transformation is a function of the fractional amount of austenite present, 1-f(T), and the rate of nucleation (per unit temperature change) per unit volume. We are taking this nucleation rate to be constant, namely the average nucleation rate <I>:

$$\langle I \rangle = \frac{\int I\,dT}{\int dT} = \frac{C}{T_S - T_F}$$

These assumptions are given in the mathematical form by

$$-\frac{dh}{dT} = [1-f(T)]\,\frac{C}{T_S - T_F} \qquad \text{(A-10)}$$

But, from equation (A-8)

$$dh(T) = \frac{T_F - T_S}{(T_F - T)^2}\,\exp\left(-\frac{T-T_S}{T_F-T}\right)dT \qquad \text{(A-11)}$$

and hence,

$$f(T)-1 = \frac{T_F - T_S}{(T_F - T)^2}\,(T_S - T_F)\,\exp\left(-\frac{T-T_S}{T_F-T}\right)\frac{1}{C} \qquad \text{(A-12)}$$

Rearranging to get

$$f(T)-1 = -\frac{1}{C}\left(\frac{T_F - T_S}{T_F - T}\right)^2 \exp\left(-\frac{T-T_S}{T_F-T}\right) \qquad \text{(A-13)}$$

In Fig. A3, f(T)-1 is plotted versus T°K, for the same values of T_F and T_S as in Fig. A2, and for C=1.4716. This C value can be looked on as a

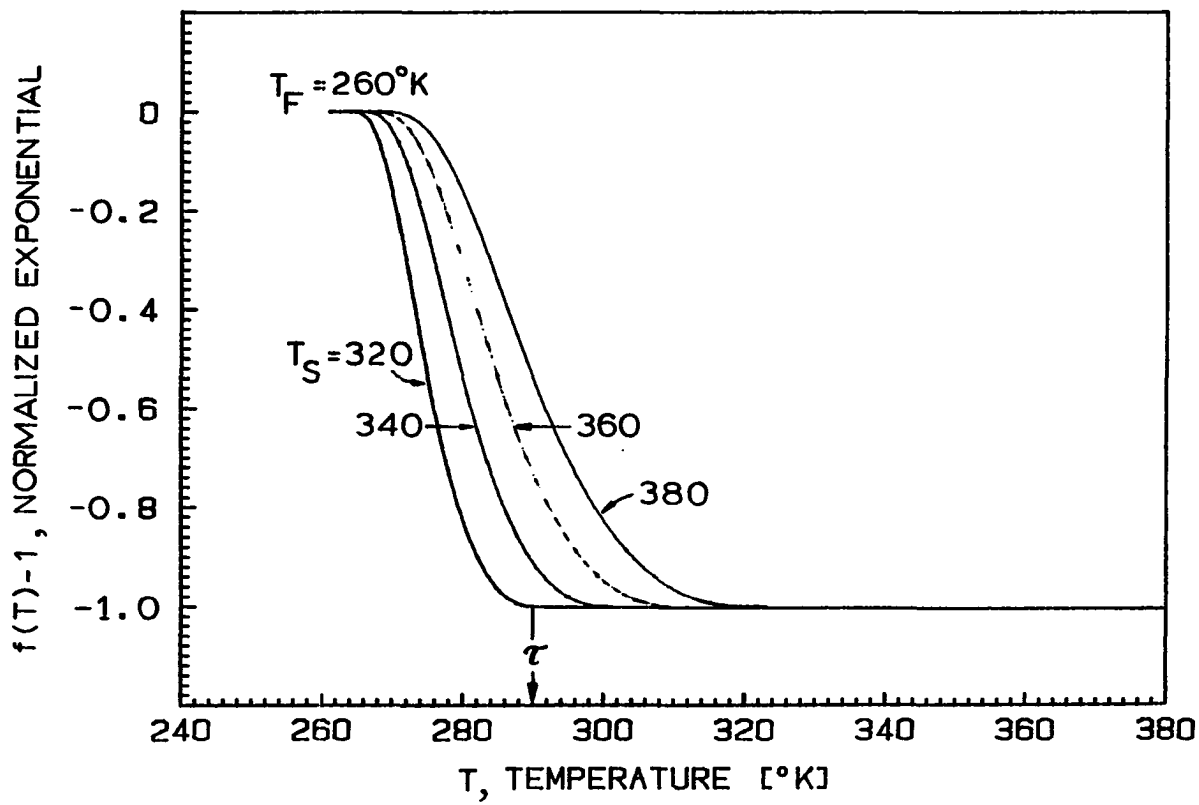


Figure A3. The dependence of the normalized exponential curves, f(T)-1 on $\Delta T=T_S-T_F$ for T_F=260°K, and ΔT=70, 80, 100, and 120°K. τ is shown for T_S=320°K only

normalization factor. Since f(T) is supposed to be a monotonic function, we define its range of validity (as given by equation (A-13)) only for $T_F^+ \leq T \leq \tau$.

Thus

$$f(T) = g(y) = 1 - (1/C)\ \left(\frac{T_F - T_S}{T_F - T}\right)^2 \exp\ \left(-\frac{T - T_S}{T_F - T}\right) \qquad \text{(A-14)}$$

This form of f(T) is analogous to the expression given by Mohamed for volume fraction of martensite [96].

This f function is plotted in Fig. A4 for the same values of T_F and T_S as in Fig. A2. To investigate f(T) of equation (A-14) further, let us define

$$(1/C)\ (T_F - T_S)^2 \equiv K^2 \qquad \text{(A-15)}$$

or

$$T_F - T_S = K\sqrt{C} \qquad \text{(A-16)}$$

Substitution from equation (A-16) for K into equation (A-14) yields

$$f(T) = 1 - \frac{K^2}{(T_F - T)^2} \exp\ \left(-\frac{T - T_S}{T_F - T}\right) \qquad \text{(A-17)}$$

Now, exp (-Γ) is expanded as power series of Γ:

$$-\exp(-\Gamma) = -1 + \frac{\Gamma}{1!} - \frac{\Gamma^2}{2!} + \frac{\Gamma^3}{3!} \cdots$$

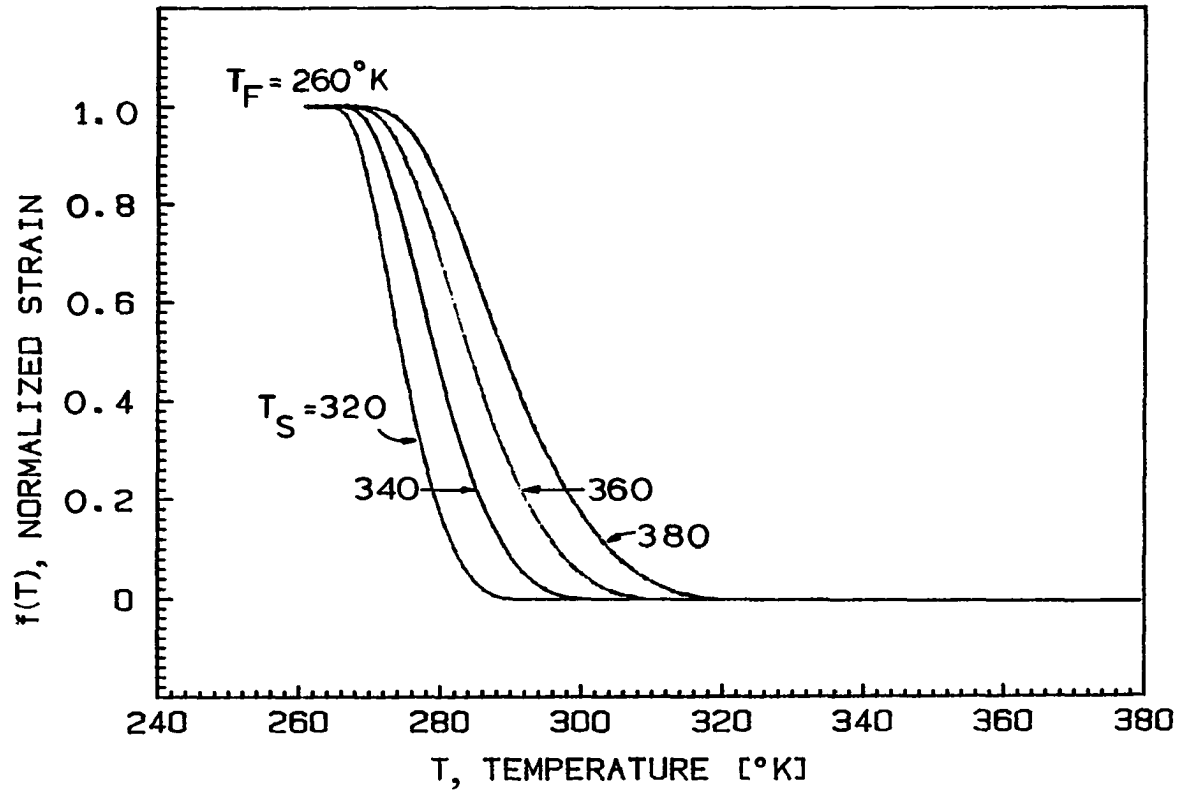


Figure A4. The dependence of the normalized strain analytical approximation curves, f(T), on $\Delta T = T_S - T_F$ for T_F=260°K, and ΔT=60, 80, 100, and 120°K

which upon substitution in equation (A-17) gives

$$f(T) = 1 - \frac{K^2}{(T_F-T)^2} + \frac{K^2(T-T_S)^2}{1!(T_F-T)^3} - \frac{K^2(T-T_S)^2}{2!(T_F-T)^4} \cdots$$

$$= \frac{(T_F-T)^2-K^2}{(T_F-T)^2} + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \frac{(T-T_s)^n}{(T_F-T)^{n+2}} K^2 \qquad \text{(A-18)}$$

with radius of convergence corresponding to the temperature range of $T_F^+ < T < \tau$.

We assume that the normalized strain is a function of temperature only (this assumption is certainly correct for a single run of a given material with defined history). Thus, f(T) is expanded as Taylor's series around T_0 with some radius of convergence δ

$$f(T) = f(T_0) + \frac{\partial f}{\partial T}\Big|_{T_0} (T-T_0) + \frac{1}{2!} \frac{\partial^2 f}{\partial T^2}\Big|_{T_0} (T-T_0)^2 + \cdots$$

$$= \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial^n f}{\partial T^n}\Big|_{T_0} (T-T_0)^n + f(T_0) \qquad \text{(A-19)}$$

Let us define:

$$T-T_S \equiv \Delta_1$$

and

$$T_0-T_S \equiv \Delta_0$$

and

$$\frac{\partial^n f}{\partial T^n}\Big|_{T=T_i} \equiv f_i^{(n)}, \quad n=1, 2,... \tag{A-20}$$

With these notations, equation (A-19) is written as

$$f(T) = f(T_o) + f_o^{(1)}(\Delta_1-\Delta_o) + \frac{1}{2!} f_o^{(2)}(\Delta_1-\Delta_o)^2+....$$

$$= \sum_{n=1}^{\infty} \frac{1}{n!} f_o^{(n)} (\Delta_1-\Delta_o)^n + f(T_o) \tag{A-21}$$

The binominal expansion of $(\Delta_1-\Delta_o)^n$ is

$$(\Delta_1-\Delta_o)^n = \sum_{j=o}^{n} (-1)^j \frac{n!}{(n-j)!j!} \Delta_1^{n-j}\Delta_o^j \tag{A-22}$$

hence, equation (A-21) is rewritten

$$f(T) = f(T_o) + \sum_{n=1}^{\infty} \frac{f_o^{(n)}}{n!} \{ \sum_{j=o}^{n} (-1)^j \frac{n!}{(n-j)!j!} \Delta_1^{n-j}\Delta_o^j \} \tag{A-23}$$

When equation (A-23) is rearranged according to powers of Δ_1, it gets the form

$$f(T) = [f(T_o)+ \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} f_o^{(j)}\Delta_o^j] + [\sum_{j=o}^{\infty} \frac{(-1)^j}{j!} f_o^{(j+1)}\Delta_o^j] \Delta_1 + ...$$

$$... + \frac{\Delta_1^n}{n!} [\sum_{j=o}^{\infty} \frac{(-1)^j}{j!} f_o^{(j+n)}\Delta_o^j] + ... \tag{A-24}$$

or

$$f(T) = \sum_{n=o}^{\infty} b_n \frac{\Delta_1^n}{n!} = b_o + \sum_{n=1}^{\infty} b_n \frac{\Delta_1^n}{n!} \tag{A-25}$$

where

$$b_o = f(T_o) + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} f_o^{(j)} \Delta_o^j$$

and

$$b_n = \sum_{j=o}^{\infty} \frac{(-1)^j}{j!} f_o^{(j+n)} \Delta_o^j \qquad n=1,2,\ldots$$

But, according to equation (A-18), f(T) is expressible also by

$$f(T) = \frac{(T_F-T)^2-K^2}{(T_F-T)^2} + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \frac{\Delta_1^n}{(T_F-T)^{n+2}} K^2 \qquad \text{(A-18)}$$

The two equations, equations (A-25) and (A-18), converge to the same limit for radius of convergence δ. Moreover, the two infinite series are equal, term by term, for equal powers of Δ_1, i.e.,

$$b_o = f(T_o) + \sum_{j=1}^{\infty} \frac{(-1)^j}{j!} f_o^{(j)} \Delta_o^j = \frac{(T_F-T)^2-K^2}{(T_F-T)^2} = 1 - \frac{K^2}{(T_F-T)^2} \qquad \text{(A-26)}$$

and

$$\frac{b_n}{n!} = \frac{1}{n!} \sum_{j=o}^{\infty} \frac{(-1)^j}{j!} f_o^{(n+j)} \Delta_o^j = \frac{(-1)^{n-1}}{n!} \frac{K^2}{(T_F-T)^{n+2}} \quad , n=1, 2\ldots \qquad \text{(A-27)}$$

Thus, the infinite set of linear equations in $f_o^{(i)}$ is

$$\frac{K^2}{(T_F-T)^2}\left\{\frac{[1-f(T_o)](T_F-T)^2}{K^2} - 1\right\} = -f_o^{(1)} \frac{\Delta_o^1}{1!} + \frac{\Delta_o^2}{2!} f_o^{(2)} - \frac{\Delta_o^3}{3!} f_o^{(3)} + \ldots$$

$$\frac{K^2}{(T_F-T)^2} = (T_F-T)\ \{f_o^{(1)} - \frac{\Delta_o^1}{1!} f_o^{(2)} + \frac{\Delta_o^2}{2!} f_o^{(3)} \ldots\}$$

$$-\frac{K^2}{(T_F-T)^2} = (T_F-T)^2\{\qquad\qquad +f_o^{(2)} - \frac{\Delta_o^1}{1!} f_o^3 + \ldots\} \qquad \text{(A-28)}$$

$$\vdots$$

or in matrix form

$$\begin{pmatrix} T_F-T & -\frac{\Delta_o^1}{1!}(T_F-T) & \frac{\Delta_o^2}{2!}(T_F-T) & \ldots & f_o^{(1)} \\ 0 & (T_F-T)^2 & -\frac{\Delta_o}{1!}(T_F-T)^2 & \ldots & f_o^{(2)} \\ 0 & 0 & (T_F-T)^3 & \ldots & f_o^{(3)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \end{pmatrix} = \frac{K^2}{(T_F-T)^2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

Together with the first equation of (A-28), we get equation (A-29):

$$\frac{K^2}{(T_F-T)^2}\left\{\frac{[1-f(T_o)](T_F-T)^2}{K^2} 1\right\} = -\frac{\Delta_o^1}{1!} f_o^{(1)} + \frac{\Delta_o^2}{2!} f_o^{(2)} - \frac{\Delta_o^3}{3!} f_o^{(3)} + \ldots$$

This infinite matrix equation, although it can be solved in principle, is not very helpful from the practical point of view. But, as we

assume from the convergence of equation (A-25), a value for n exists, say n=m, such that:

$$\left| b_m \frac{\Delta_1^m}{n!} \right| = \left| \frac{K^2 (T-T_s)^m}{(T_F-T)^{m+2}\, m!} \right| < \delta^* \qquad \text{(A-30)}$$

This equation excludes the value of $T=T_F$. With this restriction and the use of equation (A-22), it is clear that a maximum value of m=n restrains the j's to the same value, $j_{max}=n_{max}=m$. Hence, the infinite series in equation (A-24) reduces to the finite series:

$$f(T) \cong [f(T_o) + \sum_{j=1}^{m} \frac{(-1)^j}{j!} f_o^{(j)} \Delta_o^j] + \frac{\Delta_1}{1!} [\sum_{j=o}^{m-1} \frac{(-1)^j}{j!} f_o^{(j+1)} \Delta_o^j] + \ldots$$

$$+ \frac{\Delta_1^m}{m!} f_o^{(m)} \qquad \text{(A-31)}$$

which is of the form:

$$f(T) = \sum_{n=o}^{m} B_n \frac{\Delta_1^n}{n!} = B_o + \sum_{n=1}^{m} B_n \frac{\Delta_1^n}{n!} \qquad \text{(A-32)}$$

where:

$$B_o = f(T_o) + \sum_{j=1}^{m} \frac{(-1)^j}{j!} f_o^{(j)} \Delta_o^j$$

and

$$B_n = \sum_{j=o}^{m-n} \frac{(-1)^j}{j!} f_o^{(j+n)} \Delta_o^j \qquad n=1, 2, \ldots, m$$

Equation (A-18) for this finite expansion is approximated as

$$f(T) \cong \frac{(T_F-T)^2-K^2}{(T_F-T)^2} + \sum_{n=1}^{m} \frac{(-1)^{n-1}}{n!} \frac{\Delta_1^n}{(T_F-T)^{n+2}} K^2 \qquad \text{(A-33)}$$

By equating equal powers of Δ_1, we get

$$f(T_o) + \sum_{j=1}^{m} \frac{(-1)^j}{j!} f_o^{(j)} \Delta_o^j = 1 - \frac{K^2}{(T_F-T)^2} \qquad \text{(A-34)}$$

and

$$\sum_{j=o}^{m-n} \frac{(-1)^j}{j!} f_o^{(j+n)} \Delta_o^j = (-1)^{n-1} \frac{K^2}{(T_F-T)^{n+2}} \qquad n = 1, 2, \ldots, m \qquad \text{(A-35)}$$

Equation (A-35) is a set of m linear equations in the m unknowns $f_o^{(i)}$, i=1, ..., m which can easily be solved considering the triangular form of the corresponding matrix. The result is rather complicated, however, for explicit expression. Nevertheless, the bottom line here is the fact that $f_o^{(i)}$ can be expressed as a power series in Δ_o and by equation (A-31) f(T) is expressible as a power series in Δ_o, Δ_1 and T_F-T. This means that f(T) depends only on T_S, T_F, and the temperature of interest. When equation (A-35) is solved for $f_o^{(i)}$, equation (A-34) can be used to determine m, so that the process is of an iteration kind.

A similar mathematical procedure can be used for the heating part of the strain hysteresis loop. We will not, however, deal with this problem further, as the mathematical formalism which was developed so far is sufficient.

To summarize, we have found so far that

(a) The normalized strain, f(T), is expressible in terms of derivatives of f(T) with respect to temperature.

(b) The derivatives themselves can be approximated by a power series of T_F, T_S and the temperature of interest, T.

(c) The functional behavior of f(T) can be expressed as

$$f(T) = 1-(1/C)\ (\frac{T_F-T_S}{T_F-T})^2 \exp\ (-\frac{T-T_S}{T_F-T}) \qquad (A-17)$$

The functional relationship of equation (A-8) enables the definition of a temperature constant τ for the transformation, where τ is defined as the temperature at which the function h(T) is within 1/e of its final value, i.e., where $\Gamma=1$, and gives the upper limit of the validity range for f(T).

(d) In this mathematical formalism, the phase transformation property was reduced to the material constants, T_S and T_F.

7.2 Transformation Kinetics

We found in section 7.1 that:

$$f(T) = 1-(1/C)\ (\frac{T_F-T_S}{T_F-T})^2 \exp\ (-\frac{T-T_S}{T_F-T}) \qquad (A-17)$$

assuming constant rate of nucleation per unit volume. Now we assume that the temperature-rate of nucleation per unit volume I, is given by [98]:

$$I = I_o \exp\left(-\frac{\Delta G_i^*}{kT}\right) \qquad \text{(A-36)}$$

where ΔG_i^* is the change in free energy for critical i-cluster formation. Further, we assume that the "velocity" at which f(T) approaches 1, (df(T)/dT), i.e., the rate at which the martensitic transformation proceeds, depends on the nucleation rate and the relative amount of austenite retained. From these assumptions, the kinetic equation of the MPT is:

$$-\frac{df(T)}{dT} = I \cdot (1-f) \qquad \text{(A-37)}$$

or

$$\frac{df(T)}{1-f(T)} = -I\,dT \qquad \text{(A-38)}$$

On integration of both sides of equation (A-38), we get

$$-\ln(1-f(T)) = -\int I\,dT$$

From equation (A-17):

$$1-f = (1/C)\,\ln\left(\frac{T_F-T_S}{T_F-T}\right)^2 \exp\left(\frac{T-T_S}{T_F-T}\right)$$

and

$$\ln(1-f) = \ln\,(1/C) + 2\,\ln\left(\frac{T_F-T_S}{T_F-T}\right) - \frac{T-T_S}{T_F-T},$$

so that

$$\ln\,(1/C) + 2\,\ln\left(\frac{T_F-T_S}{T_F-T}\right) - \frac{T-T_S}{T_F-T} = I_o \int \exp\left(-\frac{\Delta G_i^*}{kT}\right) dT \qquad \text{(A-39)}$$

By differentiation of both sides of (A-39) with respect to T, we get

$$\frac{2}{T_F-T} - \frac{T_F-T_S}{(T_F-T)^2} = I_o \exp\left(-\frac{\Delta G_i^*}{kT}\right)$$

or

$$\frac{T_F+T_S-2T}{I_0(T_F-T)^2} = \exp\left(-\frac{\Delta G_i^*}{kT}\right)$$

From equation (A-9), $2\tau=T_F+T_S$, hence:

$$\frac{2(\tau-T)}{I_0(T_F-T)^2} = \exp\left(-\frac{\Delta G_i^*}{kT}\right) \qquad \text{(A-40)}$$

Next we solve equation (A-40) for ΔG_i^*

$$\Delta G_i^* = -kT \ln \frac{2(\tau-T)}{I_0(T_F-T)^2} \qquad \text{(A-41)}$$

Clearly, equation (A-41) has a meaning only in the range

$$T_F<T<\tau \qquad \text{(A-42)}$$

But, this is the validity range of f(T), in equation (A-14). Of course, as I_0 is unknown, ΔG_i^* can not be calculated. Nevertheless, for any reasonable values of I_0, τ, T_F, and T in the range given by equation (A-42), $\Delta G_i^*>0$, and it is a quite sharply decreasing function with decreasing T, as it should be. Thus, f(T) can be written as a function of ΔG_i^*

$$f(T) = 1-\exp\left[I_0 \int \exp\left(-\frac{\Delta G_i^*}{kT}\right) dT\right] \qquad \text{(A-43)}$$

Finally, according to Tong and Wayman [62], the dependence of M_S on applied stress is given by:

$$M_S(\sigma_a) = C_{PM}\sigma_a + M_S(0) \qquad \text{(A-45)}$$

where:

$$C_{PM} \equiv \left(\frac{\partial T}{\partial \sigma}\right) > 0$$

By analogy, let us assume that

$$T_F(\sigma_a) = T_F(0) - C_{PM}\sigma_a \qquad \text{(A-46)}$$

On substitution of these values in equation (A-41):

$$-\frac{\Delta G_i^*(0)}{kT} = \ln \frac{T_S(0)+T_F(0)-2T}{I_o(T_F(0)-T)^2} \qquad \text{(A-47)}$$

and

$$-\frac{\Delta G_i^*(\sigma_a)}{kT} = \ln \frac{T_S(0)+C_{PM}\sigma_a+T_F(0)-C_{PM}\sigma_a-2T}{I_o(T_F(0)-C_{PM}\sigma_a-T)^2}$$

or

$$-\frac{\Delta G_i^*(\sigma_a)}{kT} = \ln \frac{T_S(0)+T_F(0)-2T}{I_o(T_F(0)-C_{PM}\sigma_a-T)^2} \qquad \text{(A-48)}$$

The numerators of (A-47) and (A-48) are equal and comparison of the denominators gives:

$$\left|\frac{\Delta G_i^*(0)}{kT}\right| > \left|\frac{\Delta G_i^*(\sigma_a)}{kT}\right| \quad \text{for } T \approx \tau;\ C_{PM}\sigma_a > 2\ (\tau - T_F(0))$$

and

$$\left|\frac{\Delta G_i^*(0)}{kT}\right| > \left|\frac{\Delta G_i^*(\sigma_a)}{kT}\right| \quad \text{for } T \approx T_F(0);\ \text{any } C_{PM}\sigma_a > 0$$

But, from equations (A-36), (A-38) and (A-41):

$$\left| \frac{df(T)}{dT} \right| = I_0 \exp\left(\int I dT\right) = \frac{2(\tau-T)}{(T_F-T)^2} \exp\left(\int \frac{2(\tau-T)}{(T_F-T)^2} dT\right) \quad \text{(A-49)}$$

It follows from the above inequalities together with equations (A-47) and (A-48) that

$$\left| \frac{df(T)}{dT} \right|_{\sigma_a=0} > \left| \frac{df(T)}{dT} \right|_{\sigma_a>0} \quad \text{for } C_{PM}\sigma_a > 2(\tau-T_F(0)) \quad \text{(A-50)}$$

7.3 Comparison Between the Experimental Results and Those Predicted by the Formalism

From the experimental results, it is clear that the transformation strain behavior in a sample (like a tube) is more complicated than the strain behavior predicted by the mathematical formalism. The experimental strain depends on the direction along which the strain is measured (axial or circumferential), on the applied stress, on the sample thermomechanical history, and on its microstructure.

The applicability of equation (A-17) and its fit with experimental data points were checked for three different cases. In Table A1, the T_S, T_F and ΔT values are given for the best fitting curves in case of the sample of Fig. A1(a) and for two cases of sample P8-1, one for zero applied stress (run 2) and the other for 6 ksi applied uniaxial stress (run 9). In this table, the corresponding values of M_S, M_F, and ΔT are also given, from the axial strain versus temperature plots. In the

Table A1. Characteristic temperatures of the sample of Fig. A1 (a) under zero applied stress, and of sample P8-1 under zero and 6 ksi applied stresses, as defined by the straight line slope method and the best fit of f(T) to experimental data points

	Characteristic Temperature [K]		
	Applied Stress $\sigma = 0$ (Sample of Fig. A1(a))	Applied Stress $\sigma = 0$ (P8-1)	Applied Stress $\sigma = 6$ ksi (P8-1)
from ε_{LTP} and ε_{HTP} lines			
M_S	370	329	306
M_F	264	260	223[a]
$\Delta T = M_S - M_F$	106	69	83
from best fit of f(T) to experimental data points			
T_S	388	324	345
T_F	298	274	262
ΔT	90	50	83

[a]Extrapolated value.

case of run 9, the M_S value was found by extrapolation, as the sample did not achieve this point actually during the experiment.

In Figs. A5, A6, and A7,the experimental data points of the normalized strain (equation A-3) are plotted versus sample temperature together with the curves defined by equation (A-17). From these plots, it is evident that the fitting is best along the major steeply-rising portion of the normalized strain, while the agreement between the experimental data points and the fitting curve, equation (A-15), is poor near T_F and τ. This discrepancy can be explained partially by the inadequate method of determining the M_S and M_F temperature.

However, for values of the experimental data points in Figs. A6 and A7, it looks like $C_{pM}\sigma_a > 2(\tau - T_F(0))$, and the predicted behavior of the slopes for the curves f(T) versus T as given in equation (A-50) is clearly confirmable.

7.4 Use of the Differential Strain Hysteresis Loops for Determination of Characteristic Temperature Points

The M_S and M_F temperatures can be taken to be equal to τ and T_F, respectively, those temperatures obtained from the best fit of equation (A-17) to the experimental data points. This will not be followed, however, for two reasons:

(a) As already was pointed out, equation (A-17) gives a poor fit near the τ and T_F temperatures, and more important,

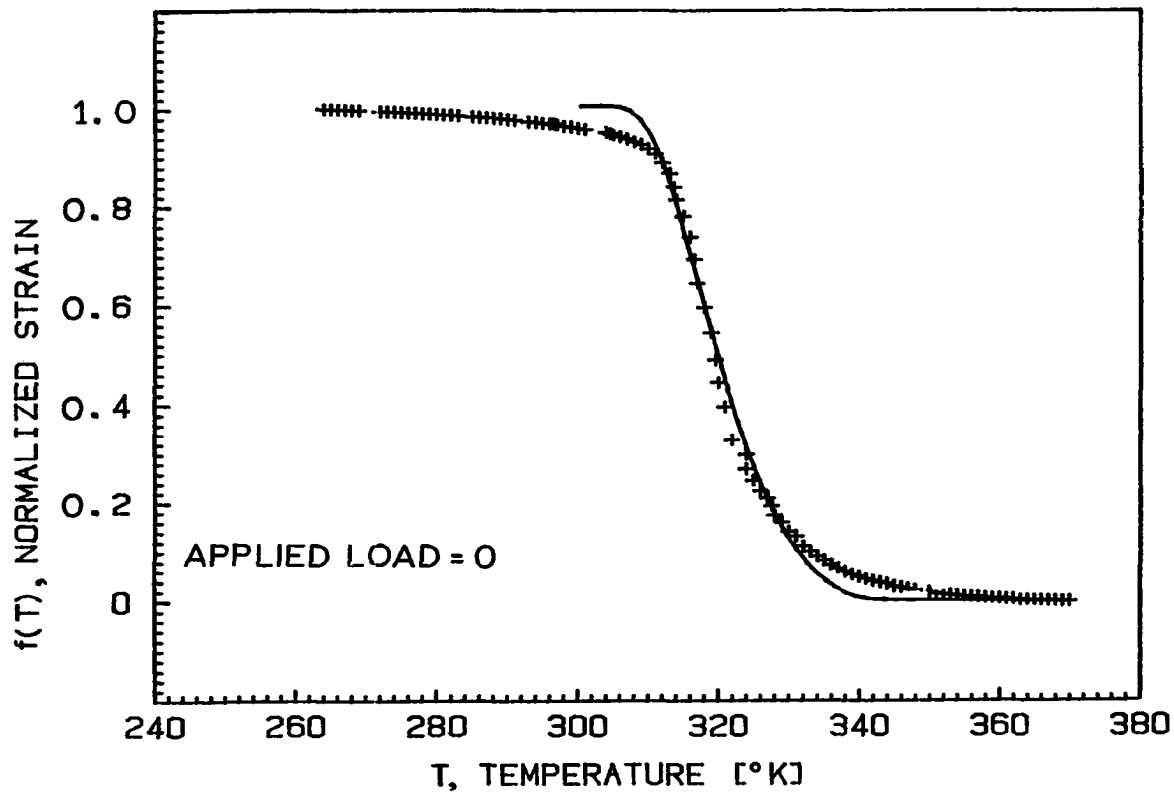


Figure A5. Fitting between normalized axial strain data points, (+), with M_S=370°K, M_F=264°K, and the analytical approximation curve, f(T), (the solid line) with T_S=388°K, T_F=298°K, for the sample of Fig. A1(a) under zero external load on cooling

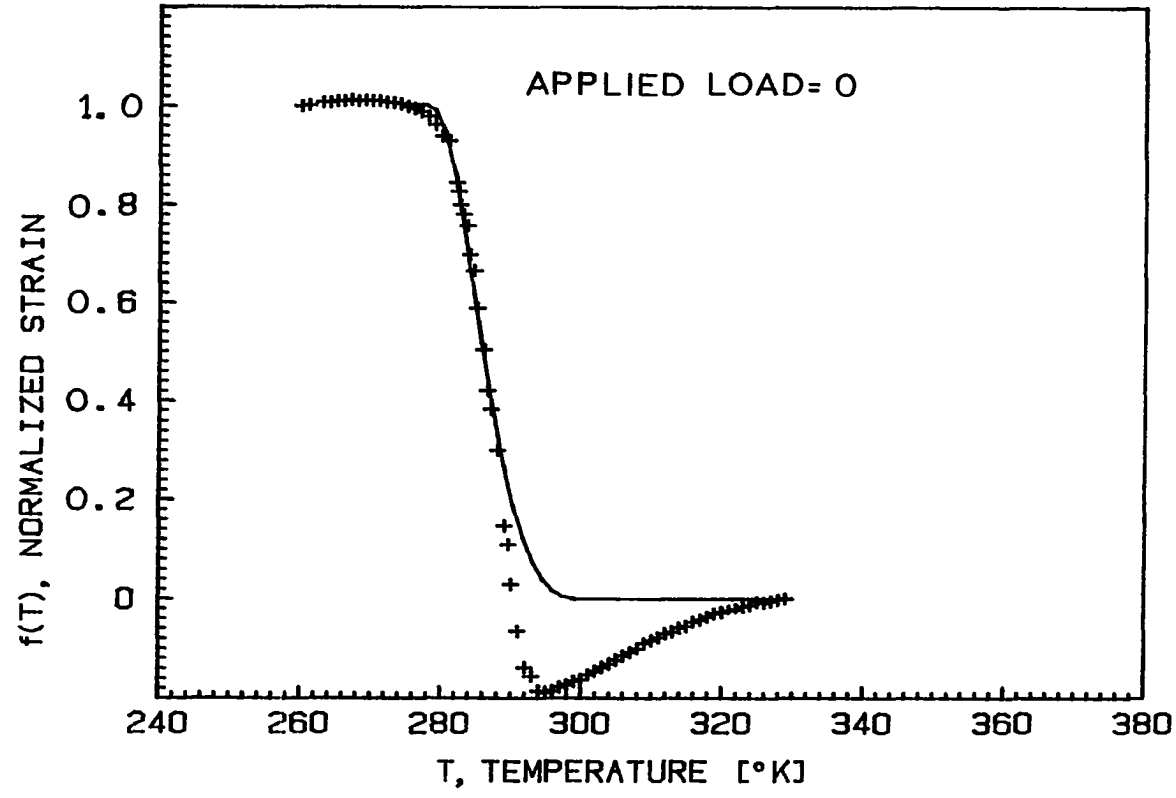


Figure A6. Fitting between normalized axial strain data points, (+), with M_S=329°K, M_F=260°K, and the analytical approximation cure, f(T), (the solid line) with T_S=324°K, T_F=274°K, for sample P8-1, under zero applied load on cooling

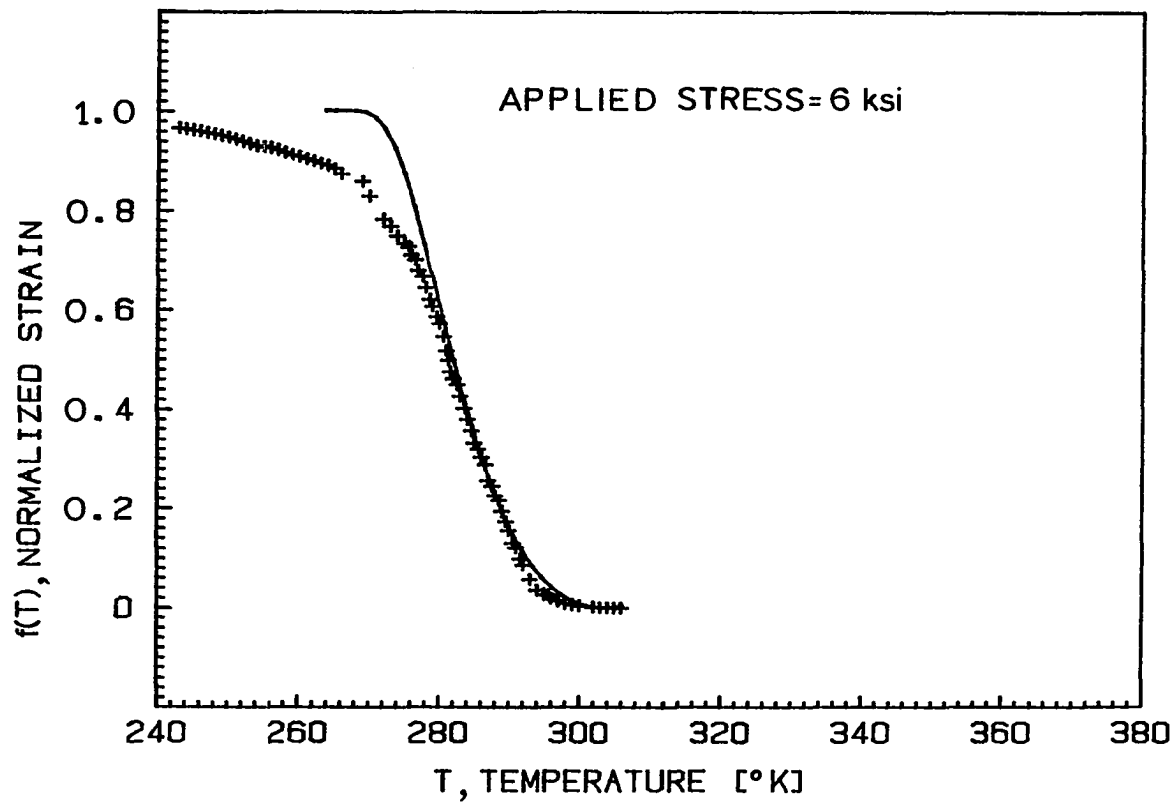


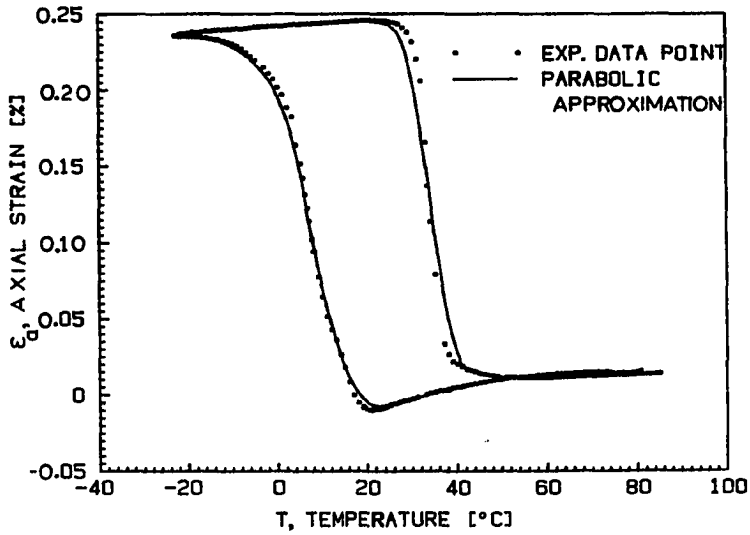
Figure A7. Fitting between normalized axial strain data points, (+), with M_S=306°K, M_F=223°K, and the analytical approximation curve, f(T), (the continuous line) with T_S=345°K, T_F=262°K, for sample P8-1, under 6 ksi applied stress on cooling

(b) We wish to define a single M_S or M_F temperature for the bulk body, independent of the direction in which the strain is measured.

To define these temperatures, we recall the fact that f(T), and hence ε(T), can be expressed as a Taylor's series of their derivatives with respect to temperature, and that these derivatives are power series of the temperature, T. These two facts enable the derivation of the derivatives of the integral strain curves by a simple computer method. First, we recognize that small deviations in the experimental data points give sharp peaks for the differential curves which is unfortunately undesirable. In order to eliminate these fluctuations, a ten point parabolic approximation was used for the integral curve. Once the parabolic function coefficients were determined, the derivative was taken for the fifth point. By elimination of the first point and addition of an eleventh point, the procedure was repeated. By this method, a relatively smooth curve was generated, which was then differentiated. Of course, approximation of the infinite expansion series by a quadratic function leaves some ambiguity for the exact shape of the integral strain. But, as shown in Figs. A8(a) and A8(b) for the axial and circumferential strains, respectively, the fit is quite good.

Since the technique of generation of differential curves is established, it is plausible to define the characteristic temperature points. A schematic transformation loop is shown in Fig. A9. The transformation

(a)



(b)

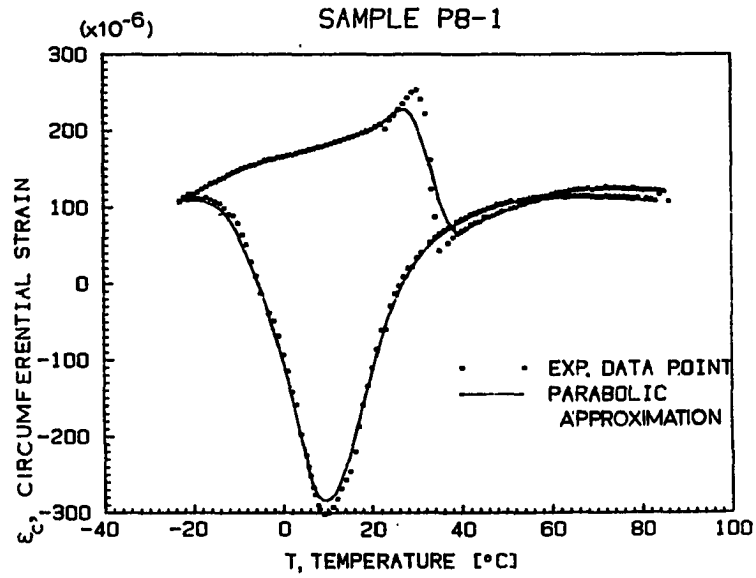


Figure A8. Strain-temperature curves. (a) Axial and (b) circumferential strain hysteresis loops of sample P8-1, on run 10 under zero load condition, (•), together with the ten point central parabolic approximation curves (solid lines)

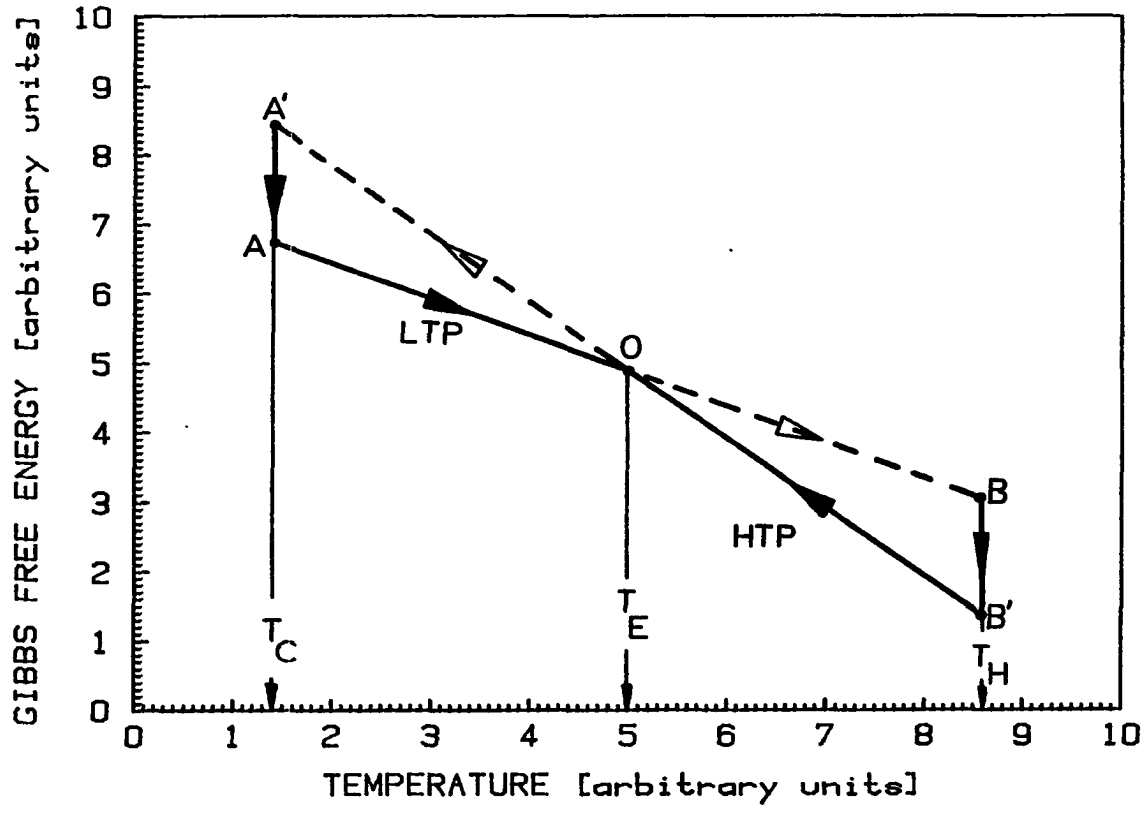


Figure A9. Schematic transformation loop in the G-T plane. Starting at A, the low temperature phase (LTP) equilibrium energy line is followed till O, which is corresponding to the equilibrium temperature. As the temperature continues to increase, the sample superheated, and at B finally phase transformation occurs. Upon cooling, the B'OA' path is followed, with subcooled material from O to A'. This plot defines three temperatures of interest, namely T_C (A'→A), T_E(O), and T_H(B→B').

loop is plotted in the ΔG-T plane, where ΔG is the change in Gibbs free energy and T is the temperature. From this loop, three temperatures of interest are defined:

T_E - the equilibrium temperature

T_C - the temperature at which the subcooled material undergoes phase transformation with a nearly infinite slope $\partial\Delta G/\partial T$

T_H - the temperature at which the superheated material transforms with nearly infinite slope $\partial\Delta G/\partial T$.

At the transformation points, $\partial\Delta G/\partial T$ is not continuous but shows the general characteristic behavior of the derivative of a physical property under first order phase transformation.

It turns out that the temperatures defined by the Gibbs' free energy concept, and M_S, M_F, A_S, and A_F can be derived directly and in a unique manner if instead of looking at the strain-temperature curves, the differential curves are used.

In Figs. A10 through A15, the differential axial and circumferential curves are shown versus sample temperature. They correspond to the integral curves shown in section 3.1.1 for the axial and the circumferential strain hysteresis loops of sample P8-1 in the as-fabricated condition (Figs. 20-25). As we are interested only in the procedure of defining the characteristic temperatures, the shape and magnitude of the peaks will not be discussed here. Rather, on each figure the T_C, T_E, and T_H are indicated, showing good agreement for values obtained from

the axial and circumferential derivative curves. T_C, and T_H were defined as the peak values of the axial differential curves on cooling and on heating, respectively. The axial curves were used as they are simpler than the circumferential curves. The temperature T_E was defined as the temperature between T_C and T_H at which the slopes of the axial strain loops on heating and cooling are equal. This definition is different from that given by Kaufman and Cohen [55].

The differential curves show peaks above and below the zero line. Thus, M_S, M_F, A_S, and A_F are defined as the temperatures where the differential curve leaves or approaches this zero line on cooling or heating, respectively. Since the characteristic temperatures are defined from the differential curves, they can be projected onto the integral strain curves to define their corresponding strain values. By this method, the characteristic temperatures of the bulk material are defined regardless of the direction in which the strain is measured, and as such, they are representative of the phase transformation for a given applied load.

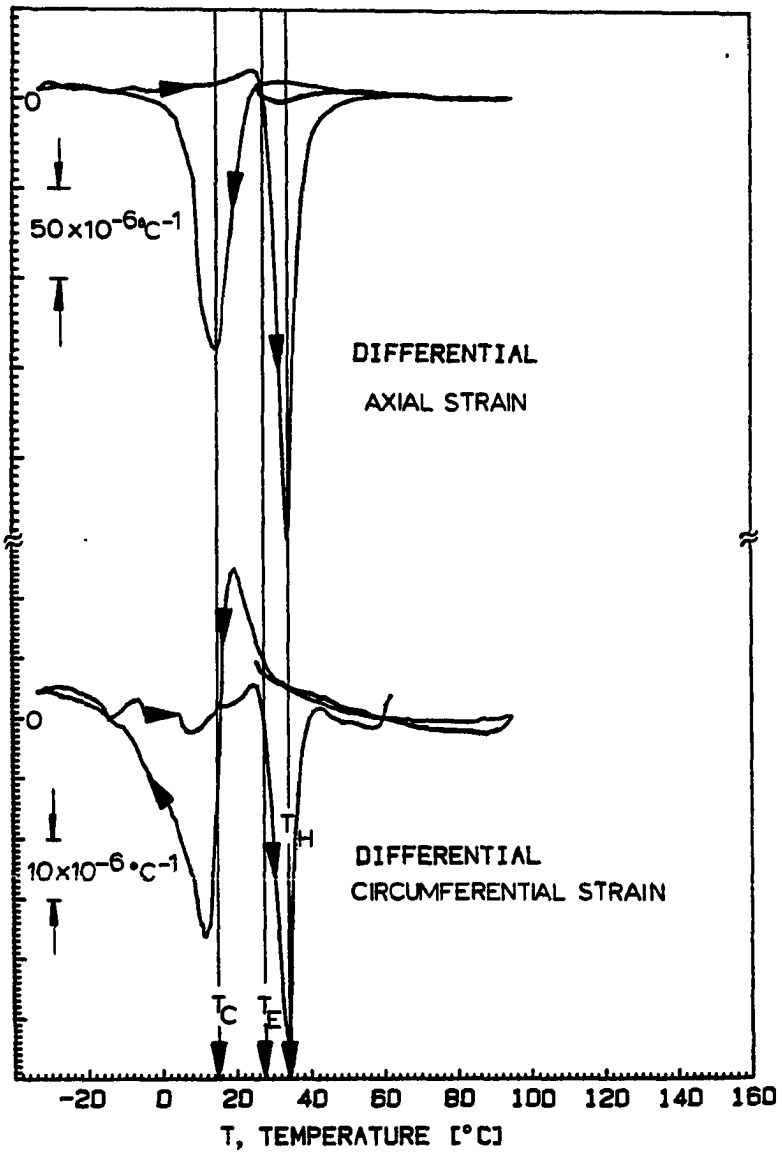


Figure A10. Differential axial and circumferential strain curves of sample P8-1 under zero load, showing T_C, T_E, and T_H of the bulk material

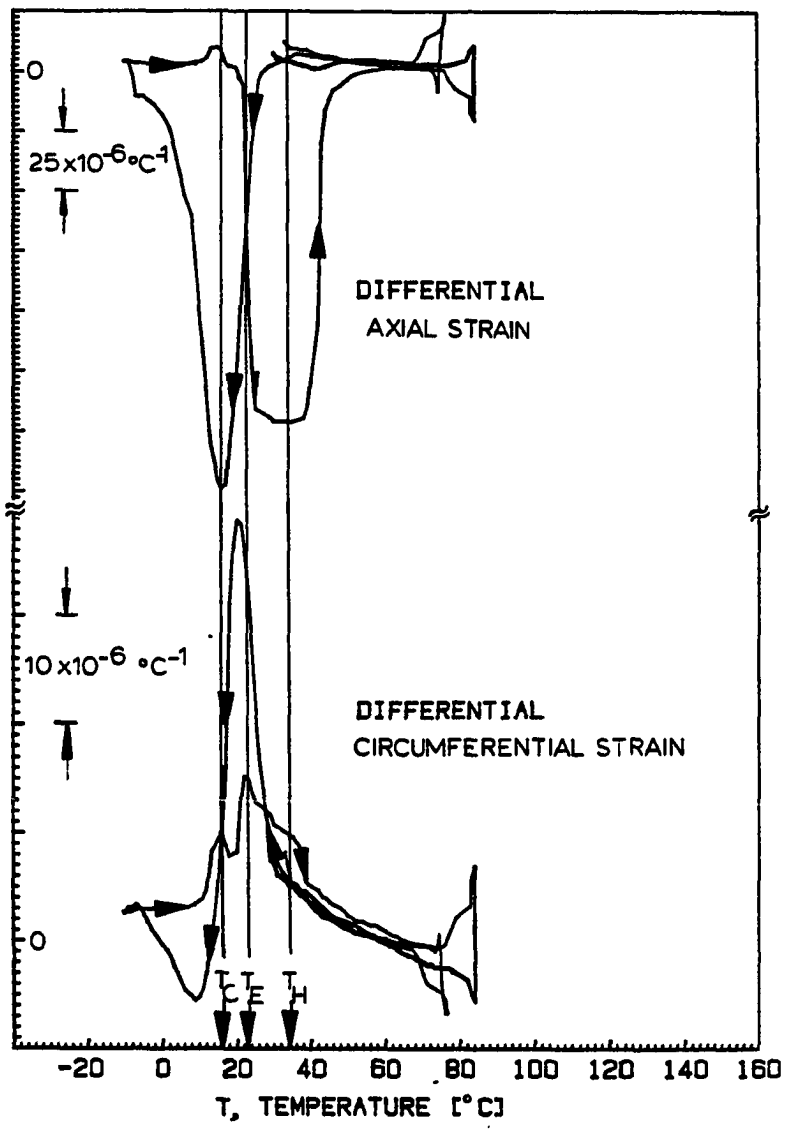


Figure A11. Differential axial and circumferential strain curves of sample P8-1 under 1.49 ksi applied stress, showing T_C, T_E, and T_H of the bulk material

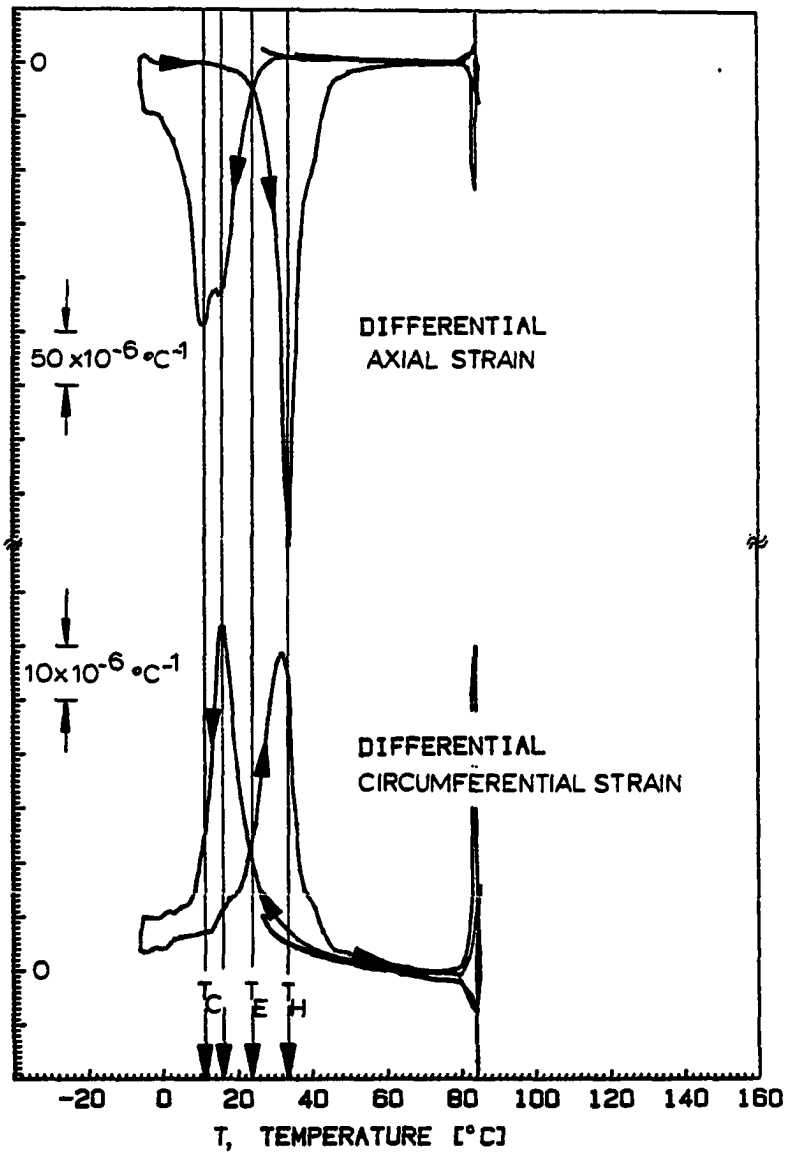


Figure A12. Differential axial and circumferential strain curves of sampel P8-1 under 3.00 ksi applied stress, showing T_C, T_E, and T_H of the bulk material.

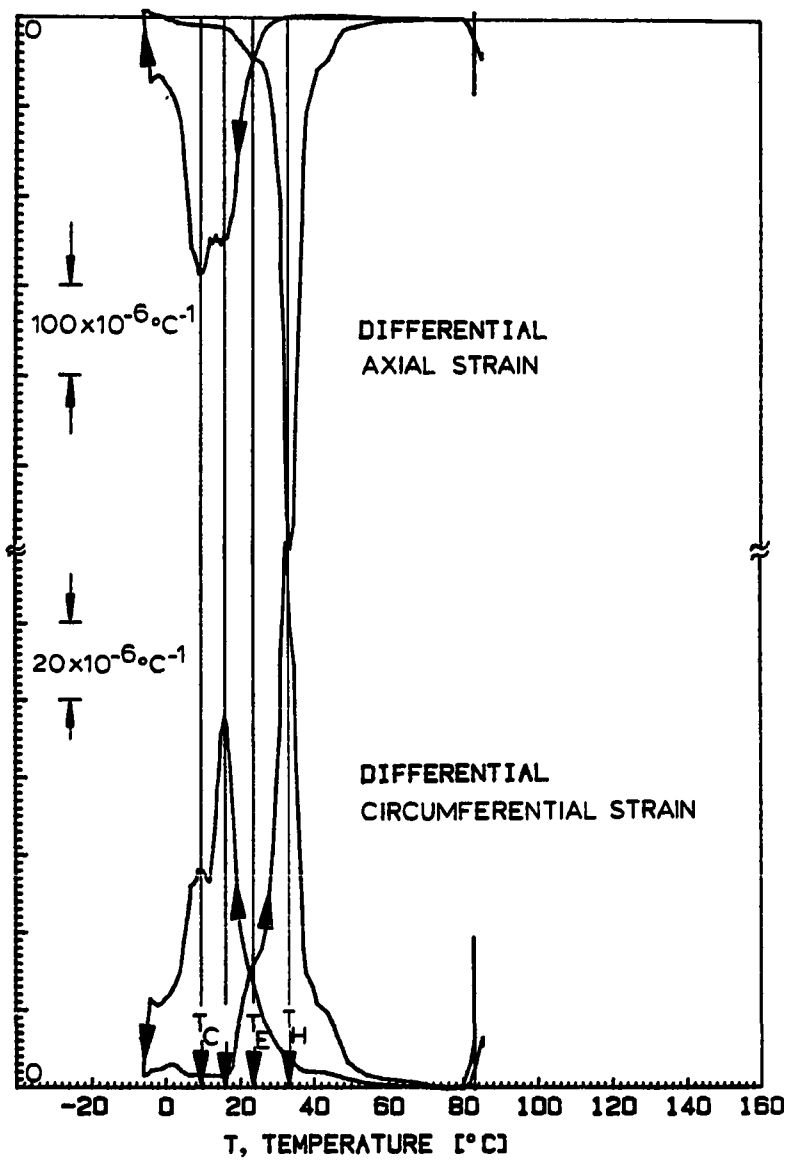


Figure A13. Diffrential axial and circumferential strain curves of sample P8-1 under 4.49 ksi applied stress, showing T_C, T_E, and T_H of the bulk material.

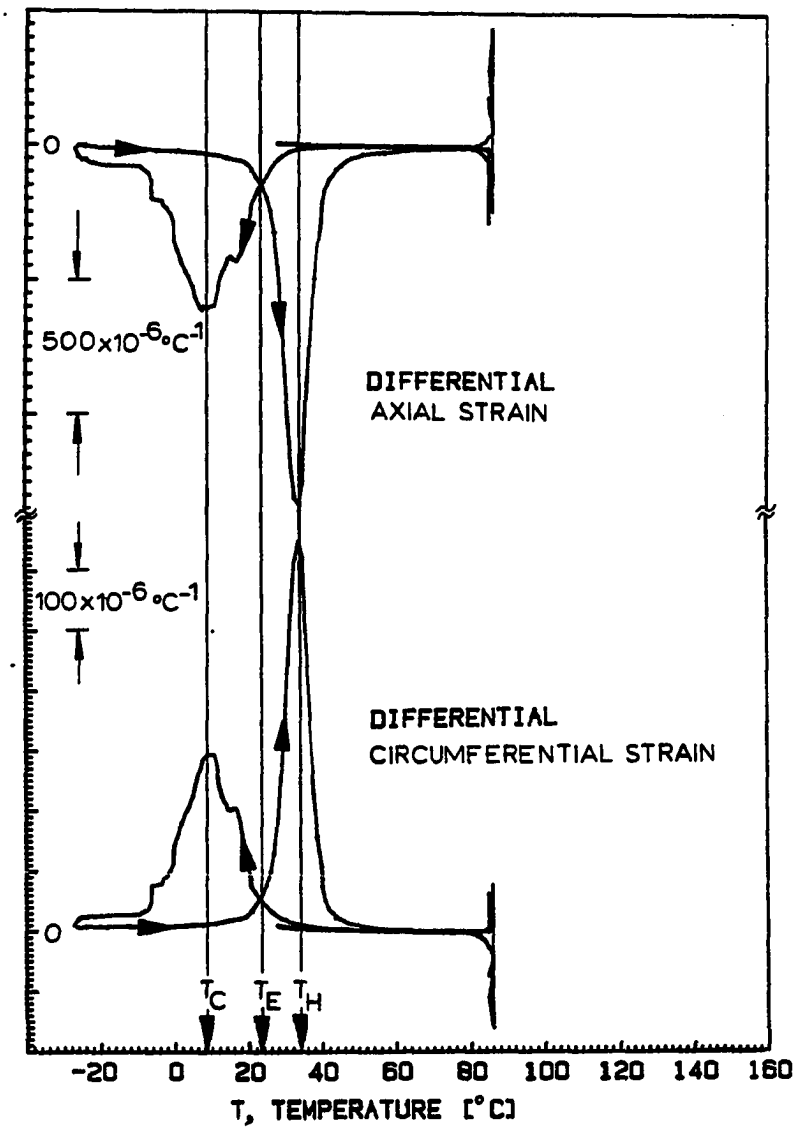


Figure A14. Differential axial and circumferential strain curves of sample P8-1 under 10.82 ksi applied stress, showing T_C, T_E, and T_H of the bulk material

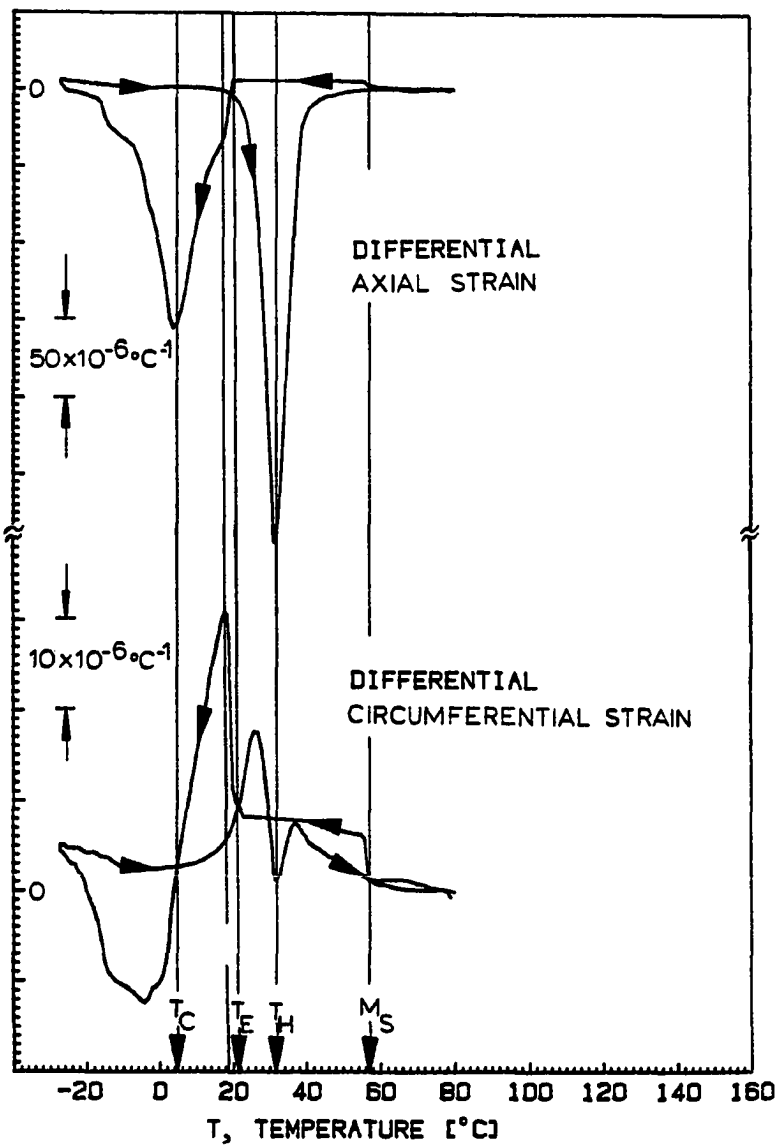


Figure A15. Differential axial and circumferential strain curves of sample P8-1 under 10.4 psi applied stress, following the run of Fig. A14, showing T_C, T_E, and T_H of the bulk material

8. APPENDIX B: APPLICATION OF THE WLR THEORY TO THE CRYSTALLOGRAPHY OF THE MPT IN NiTi

8.1 Bain Distortion and Twinning upon MPT

In the early fifties, a theory, today known as the Wechsler-Leiberman-Read (WLR) theory, was suggested to explain and predict the crystallographic features of martensitic phase transformation (MPT) [3]. It was applied successfully to MPT from the cubic to the tetragonal (InTl) [72], from the cubic to the orthorombic (AuCd) [99], and from the cubic to the monoclinic (NiTi) [100] systems. In the following, the crystallographic features of the MPT in Ni-Ti will be briefly reviewed and the WLR theory applied to calculate the total distortions.

The HTP NiTi austenite has a b.c.c. B2 type lattice structure with lattice parameter a_0=3.015Å. The Bain or lattice distortion deforms this lattice to the monoclinic (distorted orthorombic B19) lattice, with lattice parameters, a=2.889Å, b=4.120Å, c=4.622Å and $\beta=\sphericalangle(a, c)=96.8°$ [33]. In Fig. B1, the martensite (a) and the austenite (b) lattice are shown together with the axis systems $\hat{a}$, $\hat{b}$, and $\hat{c}$ (the monoclinic [M] axis system), $\hat{i}'$, $\hat{j}'$, and $\hat{k}'$ (directions [1$\bar{1}$0], [110] and [001] in the austenite, the [P'] axis system), and $\hat{i}$, $\hat{j}$, and $\hat{k}$ (cubic axes in the austenite, the [P] axis system). From an examination of these figures, it is clear that the Bain distortion matrix in [P'] (where [x] means here: the x axis system, i.e., the system in which x_1, x_2, x_3 are orthonormal) is given by [100]:

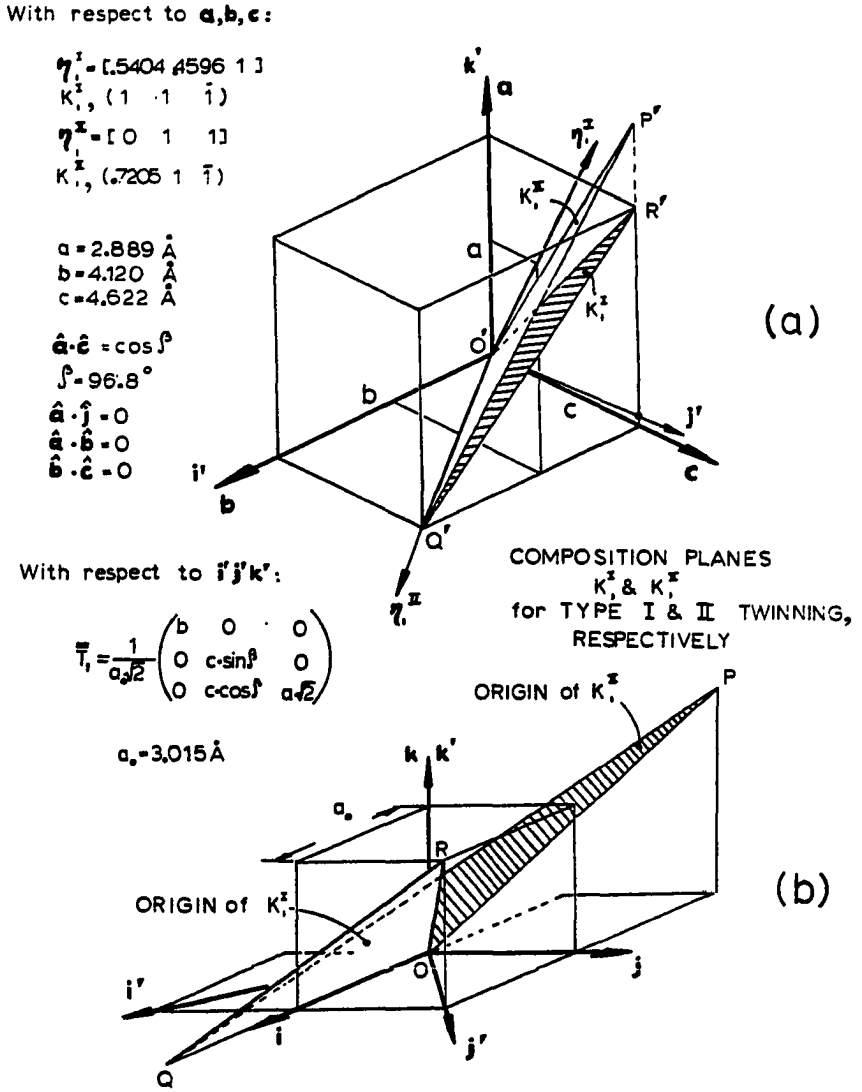


Figure B1. Lattice structure. (a) The martensite lattice structure of NiTi, constructed by shearing the orthorombic lattice along $\hat{a}$. $\hat{a}$, $\hat{b}$, $\hat{c}$ are the lattice vectors of the monoclinic system, $\hat{i}'$, $\hat{j}'$, and $\hat{k}'$ of the orthorombic. Plane O'P'Q' is the composition plane of type II twinning, and it is irrational. Plane O'R'Q' is that of the type I twinning. (b) The original austenite lattice, with $\hat{i}$, $\hat{j}$, and $\hat{k}$, being the cubic (β_1) lattice vectors. Planes OPQ and ORQ are the planes that will become the composition planes O'P'Q' and O'R'Q' in (a), respectively, by applying the Bain distortion matrix $\bar{T}_1$

$$\bar{\bar{T}}_1[P'] = \frac{1}{a_o\sqrt{2}} \begin{pmatrix} b & 0 & 0 \\ 0 & c\cdot\sin\beta & 0 \\ 0 & c\cdot\cos\beta & a\sqrt{2} \end{pmatrix} \qquad \text{(B-1)}$$

It is clear from the WLR theory that the Bain distortion by itself is unable to create an invariant plane strain (along the habit plane), and hence a lattice invariant distortion is necessary in the MPT process. This distortion can be slipping or twinning. Experimental evidence shows that the lattice invariant distortion in Ni-Ti is that of twinning, and therefore only this coherent motion of atoms will be discussed here.

Fig. B2, adopted from [101], shows the result of internal twinning of the parallelogram ABCD by shearing along the twinning plane whose trace in the plane of the paper is EF; thus, the points B and C become B' and C', respectively. The distance and direction along which these points are translated (proportional to the distance of B and C from plane EF) are given by $\vec{g}$. The distance g is smaller than the lattice parameter. It can be shown that by choosing appropriate values of the fractional volumes of the two twin related regions, 1-x and x, an invariant plane strain is achievable [3].

Formally, there are many twinning systems [102,103], and the suitable one for applying the theory of WLR must be chosen according to the experimental evidence. In NiTi, both type I and type II twinning are observed experimentally [100, 104].

The homogeneous shear due to twinning is depicted in Fig. B4 (adapted from ref. [103]). In the two dimensional twin-related crystals

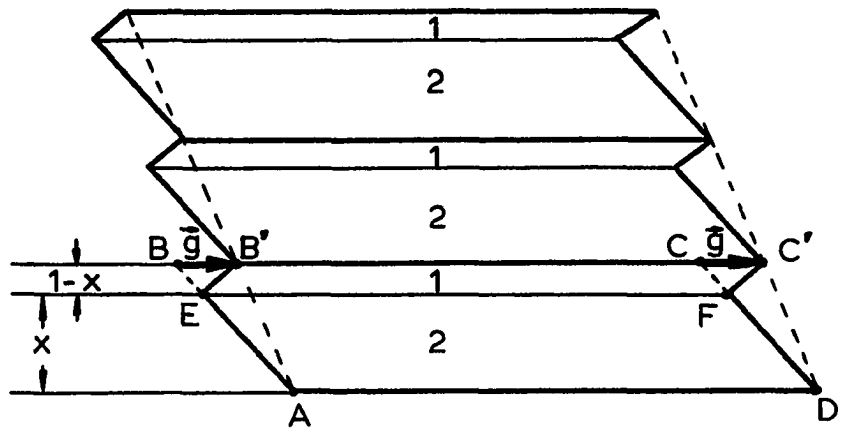


Figure B2. Production of invariant plane strain along the habit plane (not shown) by twinning the original crystal ABCD along the twinning plane whose trace is EF (perpendicular to the plane of the figure) to EB'C'F and AEFD such that their fractional volume is 1-x and x, respectively (adopted from ref. [91])

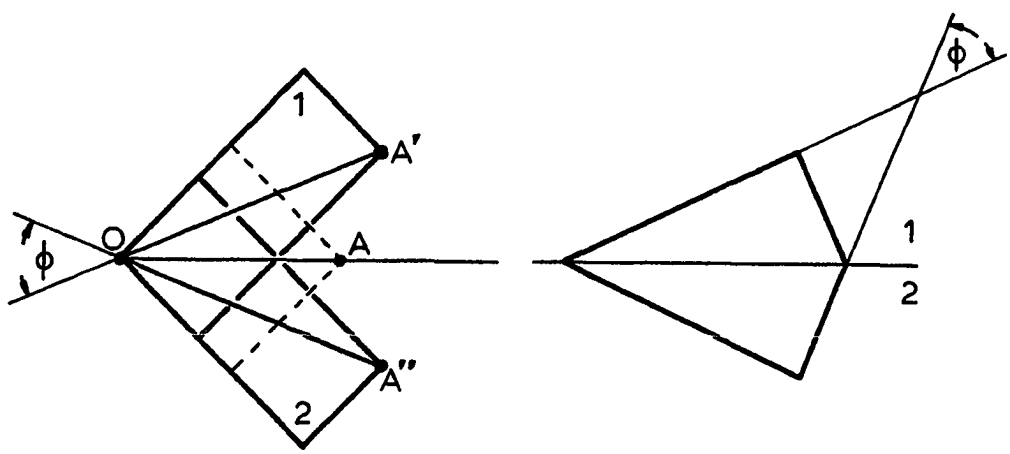
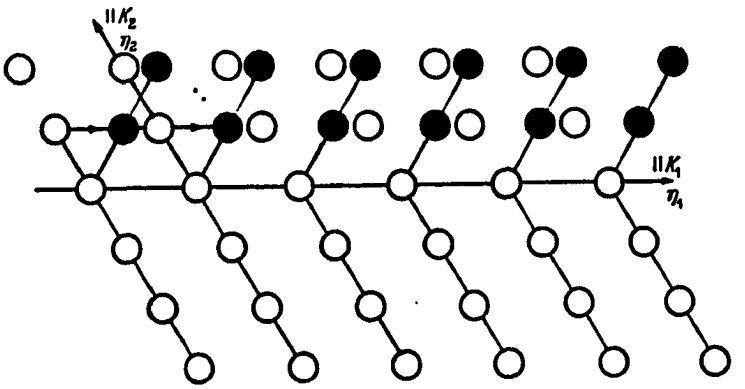


Figure B3. The rotation that produces coherent twins. (a) The original austenite (----) and two twin-related martensite regions, (1) and (2), prior to rotation along a normal to the composition plane, and (b), regions (1) and (2) after rotation, showing the geometric relation between twins (after ref. [91])

(a)



(b)

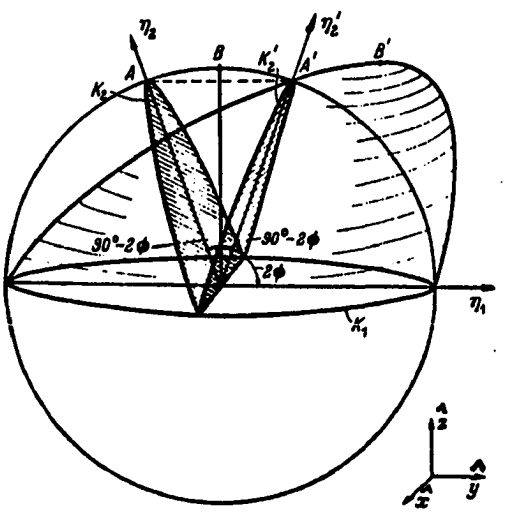


Figure B4. Representation of twinning. (a) Atom movements. (b) Deformation of a sphere by twinning to an equivalent ellipsoid by the shear $S_t=AA'=2\cot 2\Phi$. A→A', B→B'

shown in Fig. B4(a), the shear direction $\vec{\eta}_1$ (the line of intersection of the composition plane K_1 with the plane of shear) and $\vec{\eta}_2$ (the unique direction out of plane K_1 that is unchanged in length) are shown. A sphere is distorted by the twin shear into an ellipsoid of the same volume, as shown schematically in Fig. B4(b), and the related distortion matrix in the x, y, z axis system is given by:

$$\bar{\bar{S}}_t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & S_t \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{(B-2)}$$

where the amount of shear is $S_t = 2\cot 2\phi$. Clearly $\det\bar{\bar{S}}_t = 1$, indicating no change in volume due to this shear.

The twinning (sometimes called glide twinning to differentiate it from rotation twinning [105]) must be lattice invariant. Hence, three noncoplanar lattice vectors exist such that their lengths and the angles between them are not affected by the twinning, i.e., they are not distorted relative to each other. The three lattice vectors may be $\vec{\eta}_2$ and any two in K_1, in which case $\vec{\eta}_2$ and K_1 are rational and the twinning is called Type I. The three lattice vectors may also be $\vec{\eta}_1$ and any two in K_2, in which case $\vec{\eta}_1$ and K_2 are rational and the twinning is Type II [106].

Finally, as the lattice vectors in the habit plane cannot be rotated, a rotation is needed in order to satisfy all the imposed characteristics of MPT.

The total distortion matrix $\bar{\bar{E}}$ is given by [3]:

$$\bar{\bar{E}} = (1-x)\bar{\bar{M}}_1 + x\bar{\bar{M}}_2 \tag{B-3}$$

where $\bar{\bar{M}}_1$ and $\bar{\bar{M}}_2$ are the distortion matrices in twin regions 1 and 2 as illustrated in Fig. B2. $\bar{\bar{M}}_1$ and $\bar{\bar{M}}_2$ may be resolved into

$$\bar{\bar{M}}_1 = \bar{\bar{\Phi}}_1\bar{\bar{T}}_1 \tag{B-4}$$

and

$$\bar{\bar{M}}_2 = \bar{\bar{\Phi}}_2\bar{\bar{T}}_2 \tag{B-5}$$

where $\bar{\bar{T}}_1$ for twin 1 is given in equation (B-1) and $\bar{\bar{T}}_2$ for twin 2 is crystallographically equivalent to $\bar{\bar{T}}_1$ and is given below. $\bar{\bar{\Phi}}_1$ and $\bar{\bar{\Phi}}_2$ are rotations for twins 1 and 2, respectively. The relative rotation needed for coherent twin formation, as shown in Fig. B3, is given by

$$\bar{\bar{\Phi}} = \bar{\bar{\Phi}}_1^{-1}\,\bar{\bar{\Phi}}_2 \tag{B-6}$$

Equations (B-3) to (B-6) give

$$\bar{\bar{E}} = \bar{\bar{\Phi}}_1[(1-x)\bar{\bar{T}}_1 + x\bar{\bar{\Phi}}\bar{\bar{T}}_2] \tag{B-7}$$

$$= \bar{\bar{\Phi}}_1\bar{\bar{F}} \tag{B-8}$$

where we have set

$$\bar{\bar{F}} = (1-x)\bar{\bar{T}}_1 + x\bar{\bar{\Phi}}\bar{\bar{T}}_2 \tag{B-9}$$

The total shape change can be depicted as a homogeneous simple shear parallel to the habit plane and an elongation normal to it, as shown in

Fig. B5 (after refs. [3] and [101]). In this figure $\hat{z}_0$ is in the direction of the habit plane normal, and $\hat{x}_0$ is in the direction of shear. The total distortion $\bar{\bar{E}}$ in the $\hat{x}_0$, $\hat{y}_0$, $\hat{z}_0$ axis system can be written:

$$\bar{\bar{E}} = \begin{pmatrix} 1 & 0 & S \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_1\lambda_2\lambda_3 \end{pmatrix} \qquad \text{(B-10)}$$

where S is the magnitude of the projection of the displacement BB" onto the habit plane, and λ_1, λ_2, and λ_3 are the eigenvalues of the symmetric part of $\bar{\bar{E}}$, where for a plane of zero macroscopic distortion one of the λ's, say λ_3, is equal to unity. Hence, $\lambda_1\lambda_2\lambda_3$ is equal to the volume ratio.

Electron microscope observations by a number of investigators [33, 96, 100, 107] have indicated that the martensite twins in Ni-Ti correspond to Type I twinning with the K_1 plane being $\{111\}_M$, where the [M] or martensite axes are $\hat{a}$, $\hat{b}$, $\hat{c}$, as shown in Fig. B1. Furthermore, Knowles and Smith [100] applied the WLR theory based on this twinning mode and found the theoretical predictions to be in agreement with the available experimental results for the habit plane, orientation relationships, and relative amounts of the two twin orientations. Another feature of the transformation is the macroscopic distortion, which was not fully treated by Knowles and Smith [100], but is of special interest here. The purpose of this appendix is, therefore, to calculate the macroscopic distortion. Stated otherwise, the total distortion matrix $\bar{\bar{E}}$

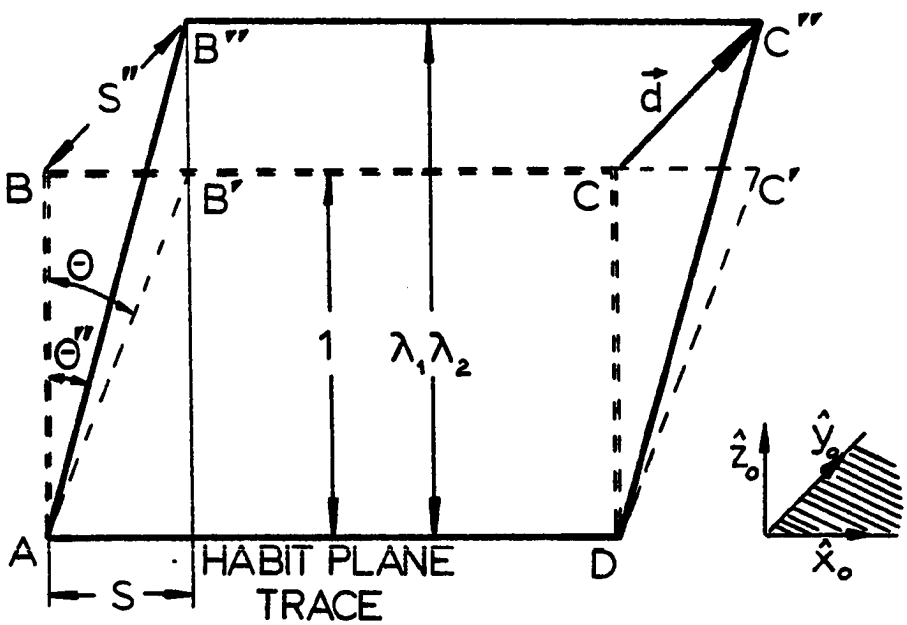


Figure B5. The total macroscopic transformation distortion is an invariant plane strain that takes ABCD in AB"C"D. It is equivalent to a simple shear with shear direction lying in the habit plane and an elongation perpendicular to it. (After Refs. 3 and 91)

in equation (B-10) transforms a sphere into an ellipsoid. In this appendix, the semi-axes of this ellipsoid are determined based on $\{111\}_M$ twinning.

8.1.1 The twin plane and its generating plane in NiTi for Type I $\{111\}_M$ twinning

In the following, $(11\bar{1})_M$ is chosen as the particular $\{111\}_M$-type plane. This plane is spanned out by the vectors O'Q' and O'R' (Fig. B1(a)), which are given by:

$$O'Q'[M] = (o,b,c)$$
$$O'R'[M] = (a,o,c)$$

where [M] indicates components with respect to the martensite axes, $\hat{a}$, $\hat{b}$, and $\hat{c}$. Another axis system is the parent cubic B2 system, $\hat{i}$, $\hat{j}$, $\hat{k}$, which are indicated by [P]. For any vector, with Miller indices [h,k,l] in [M], i.e., for any vector V[M]=(ha,kb,lc), we find:

$$\vec{V}[P] = \left(\frac{kb}{\sqrt{2}} + \frac{lc}{\sqrt{2}}\cdot\sin\beta \quad -\frac{kb}{\sqrt{2}} + \frac{lc}{\sqrt{2}}\cdot\sin\beta \quad ha + lc\cos\beta\right) \qquad (B-11)$$

Thus, we find

$$\vec{O'Q'}[P] = \left(\frac{1}{\sqrt{2}}(b+c\sin\beta) \quad \frac{1}{\sqrt{2}}(-b+c\sin\beta) \quad c\cos\beta\right) \qquad (B-12)$$

$$\vec{O'R'}[P] = \left(\frac{1}{\sqrt{2}}c\sin\beta \quad \frac{1}{\sqrt{2}}c\sin\beta \quad a+c\cos\beta\right) \qquad (B-13)$$

From equation (B-1):

$$\bar{T}_1^{-1}[P'] = a_o\begin{pmatrix} \frac{\sqrt{2}}{b} & 0 & 0 \\ 0 & \frac{1}{c}\frac{\sqrt{2}}{\sin\beta} & 0 \\ 0 & -\frac{1}{a}\frac{\cos\beta}{\sin\beta} & \frac{1}{a} \end{pmatrix} \qquad (B-14)$$

Also, the orthogonal transformation matrix from [P'] to [P] is:

$$\bar{\bar{\Theta}}\,[P/P'] = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{(B-15)}$$

and

$$\bar{\bar{\Theta}}[P/P']\bar{\bar{T}}_1^{-1}[P']\ \bar{\bar{\Theta}}^*\ [P/P'] = \bar{\bar{T}}_1^{-1}[P] \qquad \text{(B-16)}$$

where $\bar{\bar{\Theta}}^*$ here means the transpose of $\bar{\bar{\Theta}}$.

Applying $\bar{\bar{T}}_1^{-1}[P]$ of equation (B-16) to $\vec{OQ}[P]$ and $\vec{OR}[P]$ in (B-12) and (B-13), we find that the generating vectors of these vectors in the untransformed cubic phase are

$$\vec{OQ}[P] = a_o[200]$$
$$\vec{OR}[P] = a_o[111]$$

$\vec{OQ}$ and $\vec{OR}$ span out the plane $(0\bar{1}1)$ in [P], which is the plane in the parent cubic which is the generating plane of the twin plane on MPT. That this is the case is easy to see from Fig. B6, which shows the $(0\bar{1}1)$ plane (in [P]), and the as yet undistorted vectors that generate the lattice vectors of the transformed monoclinic lattice system, $\vec{a}_1$, $\vec{b}_1$, and $\vec{c}_1$. $\hat{i}$, $\hat{j}$, and $\hat{k}$ are the unit vectors in [P], and hence in the direction of the austenite, cubic β_1 lattice vectors. $\vec{a}_2$, $\vec{b}_2$, and $\vec{c}_2$ are crystallographically equivalent to $\vec{a}_1$, $\vec{b}_1$, and $\vec{c}_1$. A sufficient and necessary condition for plane $(0\bar{1}1)$ to become the twin plane relative to the two crystallographically equivalent systems, (1) and (2), is that the corresponding vector differences,

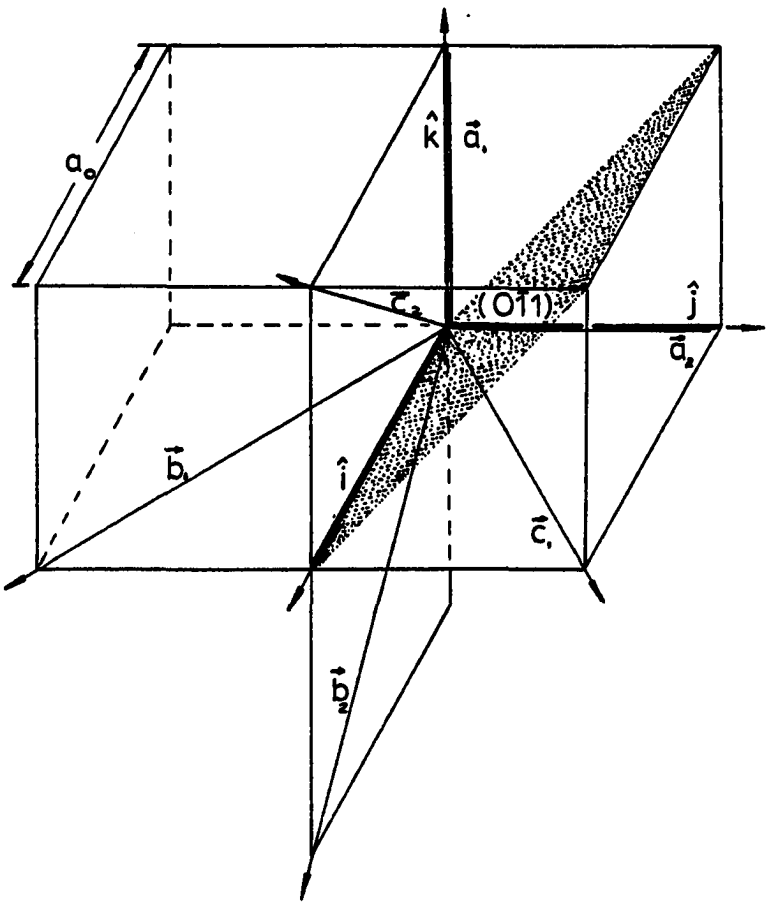


Figure B6. The unit cubic cell with lattice vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$ (i.e., the [P] axis system). Vectors $\vec{a}_1$, $\vec{b}_1$, and $\vec{c}_1$, and $\vec{a}_2$, $\vec{b}_2$ and $\vec{c}_2$ are crystallographically equivalent. Upon phase transformation they are distorted to become the lattice vectors of two monoclinic, coherent twin related lattices. The figure also shows the plane $(0\bar{1}1)_P$, the origin of the twin plane, i.e., the plane in the austenite that becomes the twin plane

$\vec{\Delta a} \equiv \vec{a}_1 - \vec{a}_2$, etc., are perpendicular to the vectors spanning out the $(0\bar{1}1)$ plane. For $\vec{\Delta c}$ the proof follows.

$$\vec{c}_2[P] = a_o\left[\frac{k}{\sqrt{2}} \quad 0 \quad \frac{k}{\sqrt{2}}\right] \text{ and } \vec{c}_1[P] = a_o\left[\frac{k}{\sqrt{2}} \quad \frac{k}{\sqrt{2}} \quad 0\right]$$

$$\therefore \vec{\Delta c} = \vec{c}_2[P] - \vec{c}_1[P] = a_o\left[0 \quad -\frac{k}{\sqrt{2}} \quad \frac{k}{\sqrt{2}}\right]$$

Let us choose the base vectors of plane $(0\bar{1}1)$ as:

$$\vec{e}_1 = [100] \text{ and } \vec{e}_2 = [111]$$

Then:

$$\vec{\Delta c} \cdot \vec{e}_1 = 0, \qquad \vec{\Delta c} \cdot \vec{e}_2 = -\frac{k}{\sqrt{2}} + \frac{k}{\sqrt{2}} = 0,$$

and similarly for $\Delta\vec{a}$ and $\Delta\vec{b}$. Hence, $(0\bar{1}1)$ is the generating plane of the twin plane for the Ni-Ti case, and also the generating plane of the twin plane for the Au-Cd case, although $\beta = 90°$ for the Au-Cd case.

8.1.2 The Bain distortions of martensite (1) and (2) for the case of Ni-Ti

From equations (B-1) and (B-9), $T_1[P]$ is given by:

$$\bar{\bar{T}}_1[P] = \frac{\sqrt{2}}{4a_o}\begin{pmatrix} c\cdot\sin\beta + b & c\cdot\sin\beta - b & 0 \\ c\cdot\sin\beta - b & c\cdot\sin\beta + b & 0 \\ \sqrt{2}\cdot\cos\beta & \sqrt{2}c\cdot\cos\beta & \frac{4a}{\sqrt{2}} \end{pmatrix} \qquad \text{(B-17)}$$

As mentioned above, there is ample experimental evidence that the twin plane in Ni-Ti is often $\{111\}_M$, which corresponds before transformation to $\{110\}_P$, or specifically in this treatment $(0\bar{1}1)_P$. Therefore, the lattice distortion for twin 2 is obtained by transforming $\bar{\bar{T}}_1[P]$ in (B-17) by reflection in $(0\bar{1}1)_P$. This gives

$$\bar{\bar{T}}_2[P] = \frac{\sqrt{2}}{4a_o} \begin{pmatrix} c\cdot\sin\beta+b & 0 & c\cdot\sin\beta-b \\ \sqrt{2}\cdot c\cdot\cos\beta & \frac{4a}{\sqrt{2}} & \sqrt{2}c\cdot\cos\beta \\ c\cdot\sin\beta-b & 0 & b+c\cdot\sin\beta \end{pmatrix} \qquad \text{(B-18)}$$

For $\beta=90°$, equation (B-18) becomes identical to equation (15) of [89] for the case of Au-Cd. (B-17) and (B-18) may be abbreviated as

$$\bar{\bar{T}}_1[P] = \frac{\sqrt{2}}{4a_o} \begin{pmatrix} u & v & 0 \\ v & u & 0 \\ w & w & z \end{pmatrix} \text{ and } \bar{\bar{T}}_2[P] = \frac{\sqrt{2}}{4a_o} \begin{pmatrix} u & 0 & v \\ w & z & w \\ v & 0 & u \end{pmatrix}$$

where: $u \equiv c\cdot\sin\beta+b$

$v \equiv c\cdot\sin\beta-b$

$w \equiv \sqrt{2}\cdot c\cdot\cos\beta$

$z \equiv 2\sqrt{2}\cdot a$

Note that only three of the above definitions are independent, as $W=(1\sqrt{2})\sqrt{4c^2-(v+u)^2}$, but of course, u, v, and z still depend on the four independent variables, a, b, c, and β.

8.2 Evaluation of the Fractional Amounts of the Twins and the Principal Distortion Values

Following the procedure outlined by Lieberman et al. for the Au-Cd case [99], we want to evaluate the fractional amounts of the two twin-related martensites. To do so, the relative rotation matrix must be known that makes the two regions coherent.

Consider two undistorted vectors, $\vec{v}_1$ and $\vec{v}_2$ in the $(0\bar{1}1)$ plane. They will become $\vec{v}'_1$ and $\vec{v}'_2$, respectively, when operated on by $\bar{\bar{T}}_1$, and $\vec{v}''_1$ and $\vec{v}''_2$ when operated on by $\bar{\bar{T}}_2$. Referring to

Fig. B3, $\vec{v}_1$ may be considered to be $\overrightarrow{OA}$, $\vec{v}_1'$ to be $\overrightarrow{OA}'$, and $\vec{v}_1''$ to be $\overrightarrow{OA}''$. But, since A' and A" are in reality the same point in the twin plane, the relative rotation, $\bar{\Phi}$ through angle ϕ must be included as indicated in equations (B-6) to (B-9).

General vectors in the austenite plane that becomes the twin plane are given by:

$$\vec{v}_1[P] = [\beta_1\ 1\ 1] \quad \text{and} \quad \vec{v}_2[P] = [\beta_2\ 1\ 1]$$

Hence:

$$\begin{aligned}
\vec{v}_1'[P] &= \bar{T}_1[P]\vec{v}_1[P] = \frac{\sqrt{2}}{4a_o}\begin{pmatrix} u\beta_1+v \\ v\beta_1+u \\ w\beta_1+w+z \end{pmatrix} \\
\vec{v}_2'[P] &= \bar{T}_1[P]\vec{v}_2[P] = \frac{\sqrt{2}}{4a_o}\begin{pmatrix} u\beta_2+v \\ v\beta_2+u \\ w\beta_2+w+z \end{pmatrix} \\
\vec{v}_1''[P] &= \bar{T}_2[P]\vec{v}_1[P] = \frac{\sqrt{2}}{4a_o}\begin{pmatrix} u\beta_1+v \\ w\beta_1+z+w \\ v\beta_1+u \end{pmatrix} \\
\vec{v}_2''[P] &= \bar{T}_2[P]\vec{v}_2[P] = \frac{\sqrt{2}}{4a_o}\begin{pmatrix} u\beta_2+v \\ w\beta_1+z+w \\ v\beta_1+u \end{pmatrix}
\end{aligned} \qquad \text{(B-19)}$$

Let us designate the unit normals by $\hat{n}'$ and $\hat{n}''$, then

$\hat{n}'[P] = \dfrac{\vec{v}_1'[P]x\vec{v}_2'[P]}{\|\vec{v}_1'[P]x\vec{v}_2'[P]\|}$, and similarly for $\hat{n}''[P]$.

Using (B-19),

$$\hat{n}'[P] = \frac{1}{\sqrt{N}}\begin{pmatrix} (u-v)w-zv \\ (u-v)w+zu \\ v^2-u^2 \end{pmatrix} \quad \hat{n}''[P] = \frac{1}{\sqrt{N}}\begin{pmatrix} zv+(v-u)w \\ u^2-v^2 \\ w(v-u)-zu \end{pmatrix} \qquad \text{(B-20)}$$

where $N=2w(u-v)^2(w+z)+z^2(u^2+v^2)+(v^2-u^2)^2$. Hence, the angle of rotation ϕ is given by

$$\cos\phi=\hat{n}'[P]\cdot\hat{n}''[P]$$

$$=(1/N)\{-[(u-v)w-zv]^2+2(u^2-v^2)[w(u-v)+zu]\} \qquad \text{(B-21)}$$

and the unit vector along the rotation axis for $\bar{\bar{\Phi}}$ is

$$\hat{u}[P] = \frac{\vec{n}''[P]x\vec{n}'[P]}{\|\vec{n}''[P]x\vec{n}'[P]\|}$$

$$\hat{u}[P] = \frac{1}{\sqrt{M}}\begin{pmatrix} -(u^2-v^2)^2 + [(w-z)u-wv]^2 \\ -w(u-v)^2(w+z)+(u^2-v^2)[(z+w)v-wu]+z^2uv \\ -w(u-v)^2(w+z)+(u^2-v^2)[(z+w)v-uw]+z^2uv \end{pmatrix} \qquad \text{(B-22)}$$

where:

$$M = \{[w(u-v)+zu]^2-(u^2-v^2)^2\}^2+$$

$$+2\{z^2uv-w(u-v)^2(w+z)+(u^2-v^2)[(z+w)v-wu]\}^2$$

For $\hat{u}[P] = (u_1, u_2, u_3)$ and ϕ, the matrix elements are given by:

$$\phi_{ij} = \delta_{ij}\cos\phi + u_iu_j(1-\cos\phi) - \epsilon_{ijk}u_k\sin\phi \qquad \text{(B-23)}$$

where

$$\delta_{ij} = \begin{cases} 0 & i\neq j \\ 1 & i\neq j \end{cases} \quad \text{and} \quad \epsilon_{ijk} = \begin{cases} 1 & \text{for cyclic permutations of ijk} \\ 0 & \text{for any two subscripts equal} \\ -1 & \text{for noncyclic permutations of ijk} \end{cases}$$

Hence, using the fact that $u_2=u_3$, $\bar{\bar{\Phi}}$ is given by equation (B-24) following

$$\bar{\bar{\Phi}}[P] = \begin{pmatrix} \cos\phi+u_1^2(1-\cos\phi) & u_1u_2(1-\cos\phi)-u_2\sin\phi & u_1u_2(1-\cos\phi)+u_2\sin\phi \\ u_2u_1(1-\cos\phi)+u_2\sin\phi & \cos\phi+u_2^2(1-\cos\phi) & u_2^2(1-\cos\phi)-u_1\sin\phi \\ u_2u_1(1-\cos\phi)-u_2\sin\phi & u_2^2(1-\cos\phi)+u_1\cos\phi & \cos\phi+u_2^2(1-\cos\phi) \end{pmatrix}$$

Now, as was shown by Lieberman et al. [99] for the Au-Cd case, if d is the principal axes system for the symmetric part of $\bar{\bar{F}}$ in equation (B-9), then for invariant plane strain to exist

$$\det(\bar{\bar{F}}_s^{*}[d]\bar{\bar{F}}_s[d]-\bar{\bar{I}}) = 0 \qquad \text{(B-24)}$$

must hold, where $\bar{\bar{F}}_S$ is the symmetric part of $\bar{\bar{F}}$, i.e. $\bar{\bar{F}}=\bar{\bar{\Psi}}\bar{\bar{F}}_S$, and $\bar{\bar{\Psi}}$ is a rotation. Also, it follows that

$$\det(\bar{\bar{F}}^{*}[P]\cdot\bar{\bar{F}}[P]) = \det(\bar{\bar{F}}_s^{*}[d]\bar{\bar{F}}_s[d])$$

so that (B-24) can be written as

$$\det(\bar{\bar{F}}^{*}[P]\bar{\bar{F}}[P]-\bar{\bar{I}})=0 \qquad \text{(B-25)}$$

from which x can be determined.

From equations (B-7) and (B-8):

$$\bar{\bar{G}}[P]\equiv\bar{\bar{F}}^{*}[P]\bar{\bar{F}}[P]=((1-x)\bar{\bar{T}}_1^{*}[P]+x\bar{\bar{T}}_2^{*}[P]\bar{\bar{\Phi}}^{*}[P])((1-x)\bar{\bar{T}}_1[P]+x\bar{\bar{\Phi}}[P]\bar{\bar{T}}_2[P])$$

Expanding in powers of x,

$$\begin{aligned}\bar{\bar{G}}-\bar{\bar{I}} \equiv\ & x^2[\bar{\bar{T}}_2^{*}\bar{\bar{T}}_2-(\bar{\bar{T}}_2^{*}\bar{\bar{\Phi}}^{*}\bar{\bar{T}}_1+\bar{\bar{T}}_1^{*}\bar{\bar{\Phi}}\bar{\bar{T}}_2)+\bar{\bar{T}}_1^{*}\bar{\bar{T}}_1] \\ & + x\,[(\bar{\bar{T}}_2^{*}\bar{\bar{\Phi}}^{*}\bar{\bar{T}}_1+\bar{\bar{T}}_1^{*}\bar{\bar{\Phi}}\bar{\bar{T}}_2)-2\bar{\bar{T}}_1^{*}\bar{\bar{T}}_1] \\ & + [\bar{\bar{T}}_1^{*}\bar{\bar{T}}_1-\bar{\bar{I}}]\end{aligned} \qquad \text{(B-26)}$$

As for any matrix $\Sigma\bar{\bar{A}}_i^{*}\bar{\bar{A}}_i$ is symmetric, it is easy to see that $\bar{\bar{G}}-\bar{\bar{I}}$ is symmetric, and moreover, each coefficient of the above power series in x is symmetric.

By solving the equation $\det(\bar{\bar{G}}-\bar{\bar{I}})=0$, x is found, and thus $\bar{\bar{G}}$ is determined. The characteristic equation of $\bar{\bar{G}}$ is given by:

$$\det(\bar{\bar{G}}-\mu\bar{\bar{I}})=0 \tag{B-27}$$

Solving the above equation and remembering that $\bar{\bar{G}}$ is symmetric, we find that the power expansion of μ takes the form:

$$\mu^3+\mu^2(-g_{11}-g_{33}-g_{22}) + \mu(g_{22}g_{33}+g_{22}g_{11}+g_{33}g_{11}-g_{13}^2-g_{12}^2-g_{23}^2)$$
$$+(g_{23}^2g_{11}+g_{12}^2g_{33}+g_{13}^2g_{22} - 2g_{23}g_{13}g_{12}-g_{11}g_{22}g_{33})=0 \tag{B-28}$$

or: $\mu^3+\mu^2a_1+\mu a_2+a_3=0$.

But if μ_i are the solutions of the characteristic equation (B-27), then

$$\prod_{i=1}^{3}(\mu-\mu_i) = 0 \tag{B-29}$$

By comparing coefficients of (B-28) and (B-29),

$$\Sigma\mu_i = -a_1$$
$$\Pi\mu_i = (-1)^3a_3$$

Since $\mu_3=1$, these equations can be solved easily for μ_1 and μ_2.

Since $\bar{\bar{G}}\equiv\bar{\bar{F}}^*\bar{\bar{F}}$, the eigenvalues of the symmetric part of $\bar{\bar{F}}$ are the square roots of the μ_i's.

Thus, in principle, x and λ_i are known. In practice, however, it is quite difficult to solve the above equations algebraically. Also, we are interested in the numerical values of λ_i in order to compare them with the experimental values. For this reason, the numerical values of

ϕ, $\bar{\bar{\Phi}}$, x, and λ_i are given below, calculated from the lattice parameters given by Otsuka et al. [33] as given at the beginning of this appendix:

$$\phi = 14.66°$$

$$\bar{\bar{\Phi}}[P] = \begin{pmatrix} 0.9792 & -0.1542 & 0.1320 \\ 0.1320 & 0.9779 & 0.1623 \\ -0.1542 & -0.1415 & 0.9779 \end{pmatrix}$$

$$x = 0.3200575$$

$$\lambda_1 = 1.054 \qquad \lambda_2 = 0.945$$

The above value of x is in good agreement with that calculated by Knowles and Smith [100], and found experimentally by Gupta and Johnson [107] the fractional deviation from the value given in [100] being 0.01‰. The deviation of $\lambda_1\lambda_2$ from the direct calculation of the volume ratio as given by $(abc \cdot \sin\beta)/(2a_0^3)$ is negligible.

Finally, the semiaxes of the total distortion ellipsoid were calculated, by using the condition that one of the principal distortions is unity. Hence, the problem is reduced from three to two dimensions, i.e., to finding the semiaxes of an ellipse and not of an ellipsoid. The two dimensional shear matrix with dilation Δ is

$$\bar{\bar{S}} = \begin{pmatrix} 1 & s \\ 0 & 1+\Delta \end{pmatrix}$$

where s and Δ are defined by (see Fig. B5):

$$1+\Delta = \lambda_1\lambda_2, \qquad s = \sqrt{(\lambda_1-1)(1-\lambda_2)}$$

Hence, if $\vec{r}$ and $\vec{r}'$ are undistorted and distorted vectors, respectively,

then: $\vec{r}'=\bar{\bar{S}}\vec{r}$. The equation of the initial undistorted unit circle is

$$\vec{r}*\vec{r} = 1$$

whereby

$$(\vec{r}')*\bar{\bar{S}}*\bar{\bar{S}}\vec{r}' = 1$$

is the equation of the distortion ellipse. In polar coordinates, $\vec{r}'$ is a function of the polar angle θ, and the major and minor semiaxes are given by

$$\frac{dr'}{d\theta} = 0$$

which gives the values

$$r_{max} = 1.054$$

$$r_{min} = 0.945$$

which agree with (B-30).

9. APPENDIX C: AN OUTLINE OF THE PREFERRED ACTIVATION OF VARIANTS MODEL (PAM)

The driving force behind the MPT is the fact that the martensite lattice is energetically more favorable than the austenite lattice, below some temperature and/or above some stress level. Hence, the free energy balance is the governing factor that determines the material behavior on phase transformation. It follows from the relatively large number (24) of crystallographically equivalent martensite variants (twice the number of slip systems in fcc metals for example) that, if there were no other consideration, the variants should be activated randomly. Thus, at the onset of the MPT in a virgin material with no applied stress, the activation of each of the 24 variants is equally likely where the transformation can be depicted as the formation of many isolated martensite islands in the surrounding austenite. As the MPT proceeds, however, other factors begin to contribute to the free energy balance. The very existence of the martensite plates introduces internal stresses, and hence changes the energy balance. As a consequence, those variants will be activated that are most capable of reducing these stresses or the related strain energy and the activation of the variants becomes less and less random. From this schematic description, it follows that in the advanced stages of MPT the variants with strain ellipsoids most closely aligned with the internal stress directions will be activated. Still, as the stress is a point function and it changes from location to location in the sample,

the average effect (integrated over the volume of the sample) of locally preferred variants in absence of external applied stresses will be small. In this first run, the macroscopic transformation strain is anticipated and found to be isotropic in sense and magnitude in all directions. This was observed for the first run of the annealed sample P8-1 under zero load, as shown in Fig. C1.

For the forward transformation under zero applied stress, the MPT begins at M_S as a random process, as far as the variants are concerned, and becomes less and less random as M_F is approached. On reverse transformation, the sequence of events must be reversed, i.e., the reversion of the martensite to austenite is quite predetermined. But, the material will not transform back exactly to the same state as it was prior to the first run.

If a given variant was active for the first run, it will more likely be activated for the second run, especially at the start of the MPT. One possibility is that after the first run, a pattern of residual stresses remains that promotes the operation of the same variant at a given location during the second transformation on cooling. Hsu [80] has reported increases in yield stress in the HTP upon transformation fatigue cycling of Ni-Ti wires, which is consistent with the introduction of residual stresses upon cycling. The residual stresses thus act as a code imprinted in the material. Following a large number of cycles, the natural selection of the remembered variants dictated by the above-

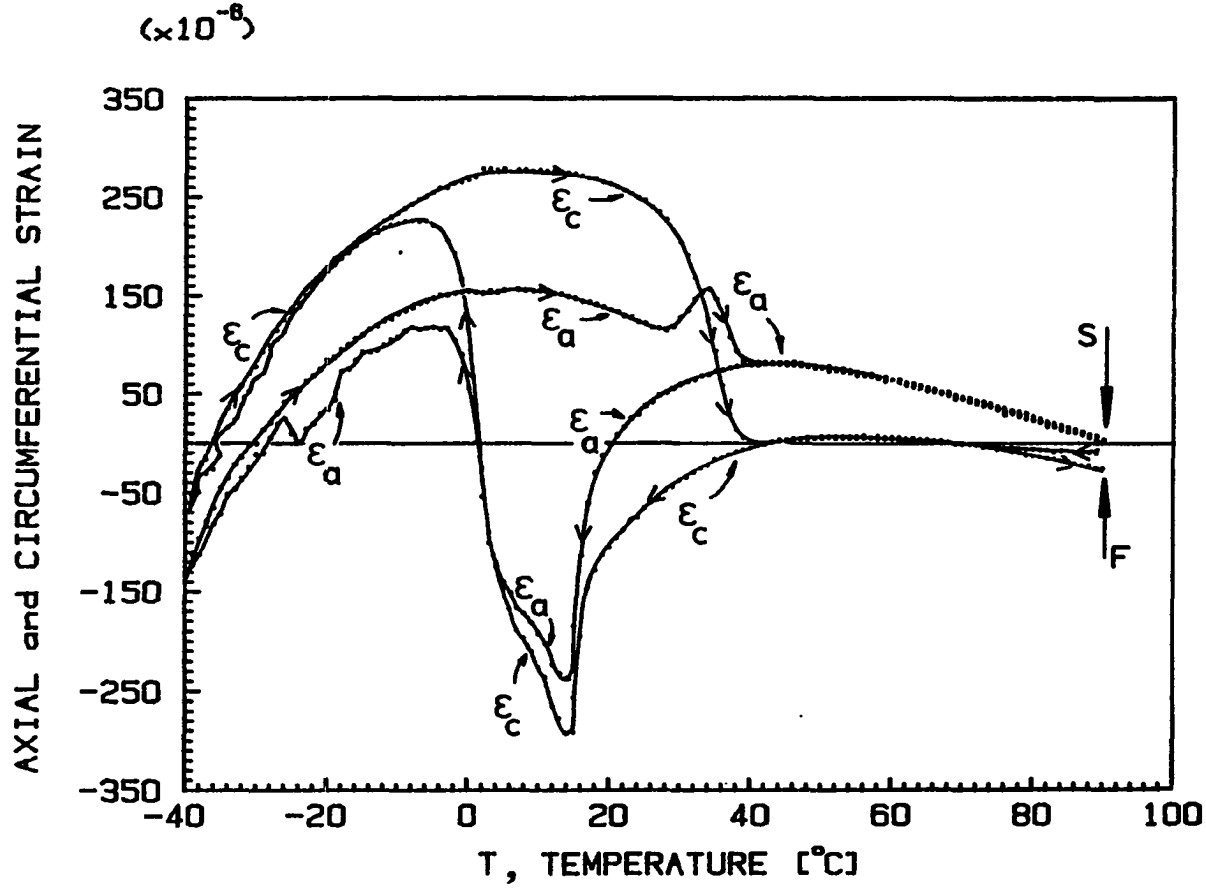


Figure C1. Axial (ε_a) and circumferential (ε_c) strains for the initial run after annealing (1 hr at 600°C). Zero applied stress

mentioned code will result in a predetermined transformation behavior. The material is then trained.

The situation is quite different when the material is thermal cycled under applied load. For simplicity, let us assume a uniaxial tensile load. When stress is applied to the HTP of an untrained material, the applied stress introduces well-defined directions that did not exist in the zero-load case. Hence, the first variants to be activated will be those that tend to relieve the applied stresses, thus lowering the free energy of the system not only by virtue of phase transformation, but also by the additional strain energy release. Clearly, the most potent variant to do so will be the one whose distortion ellipsoid has its major axis along the uniaxial tensile direction. This mechanism is depicted schematically in Fig. C2. Activation of these variants will cease only when the internal transformation strain becomes high enough compared to the strain induced by the applied stress. Variants still will be chosen deterministically, but their deformation ellipsoids will no longer be aligned along the uniaxial applied load direction. The essential point here is, that the MPT is deterministic from the very beginning, and hence, if nothing else changed, we expect an identical second run behavior. Thus, the sample is trained in one run under the influence of applied stress, instead of in many runs as is the case for not loaded runs.

Now let us assume that the applied stress is increased from run to run. Then, on the second run the most highly preferred variants will be

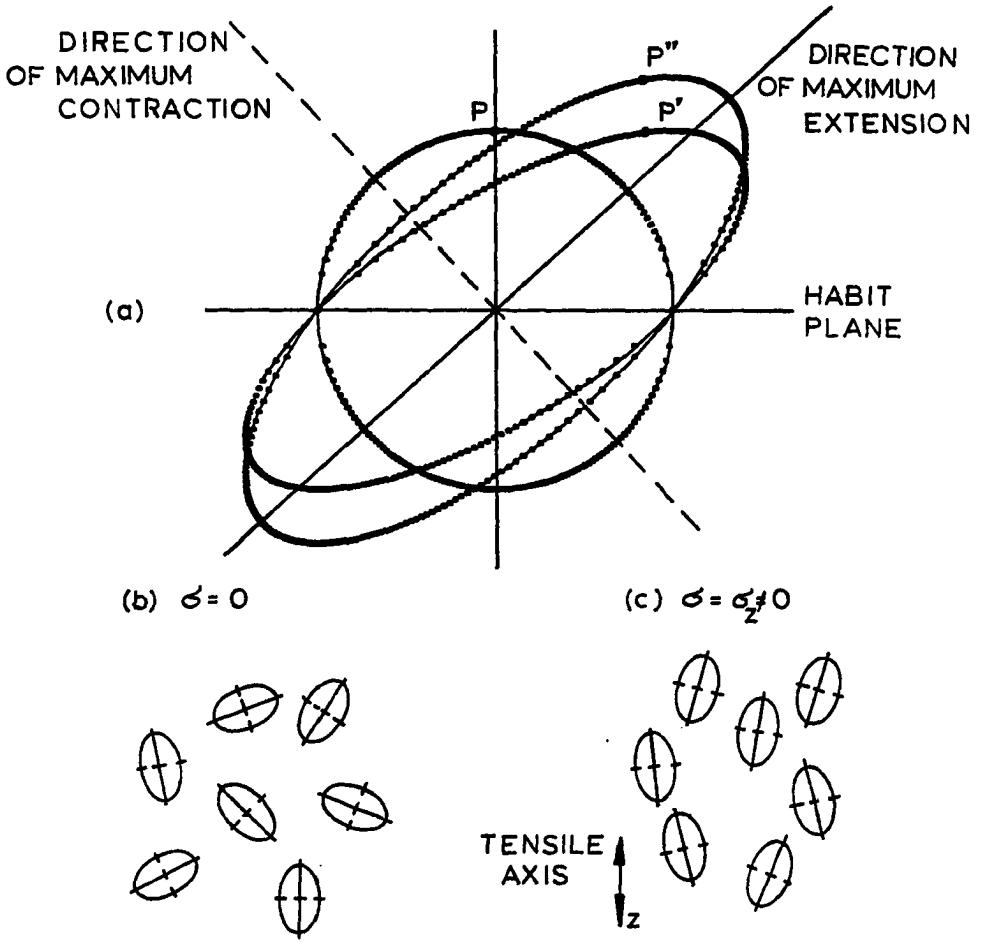


Figure C2. Schematic representation illustrating the selection of martensite variants in response to the applied stress. Figure (a) shows the shape change corresponding to an invariant plane strain. A spherical volume containing point P is transformed by a simple shear into an ellipsoid containing point P'. Then a uniaxial extension occurs perpendicular to the shear plane that gives a larger ellipsoid containing point P" to account for the transformation volume change. In Figure (b) for $\sigma=0$ all the crystallographically equivalent variants are operating, whereas in Figure (c) for $\sigma>0$ those variants are preferred for which the maximum extension direction is near the tensile axis

activated immediately, and on decrease in temperature a second less preferred variant (major axis of distortion ellipsoid less well-aligned along the applied stress axis) will be activated, as was the case for the first run. But since the applied stress is higher, the strain due to the most preferred variant will be greater before the less preferred variant operates. By this mechanism, the axial transformation strain will increase significantly with increasing applied stress, as shown in Fig. C3.

Under uniaxial tensile load, the axial direction is the major principal direction, and the other two principal directions lie in the transverse plane. Since the system is axially symmetric, the two transverse principal strains are equal. Furthermore, since the transformation volume change is small (only about 0.2%), each of the two transverse principal strains should be opposite in sign to the axial strain and one-half the magnitude. Thus, as shown in Figs. C3 and C4, the circumferential strains are negative and about one-half the magnitude of the axial strains.

On completion of several runs with ever increasing load, the sample is highly trained, and because of the previously mentioned code, it is anticipated that on reversing the loading sequence (ever decreasing tensile loads) the transformation strains will be higher than in the case of the respective increasing loading pattern. Now, if the reversed loading pattern is followed by compressive applied stress, the previous training must be washed out from the material memory. This will result at the few

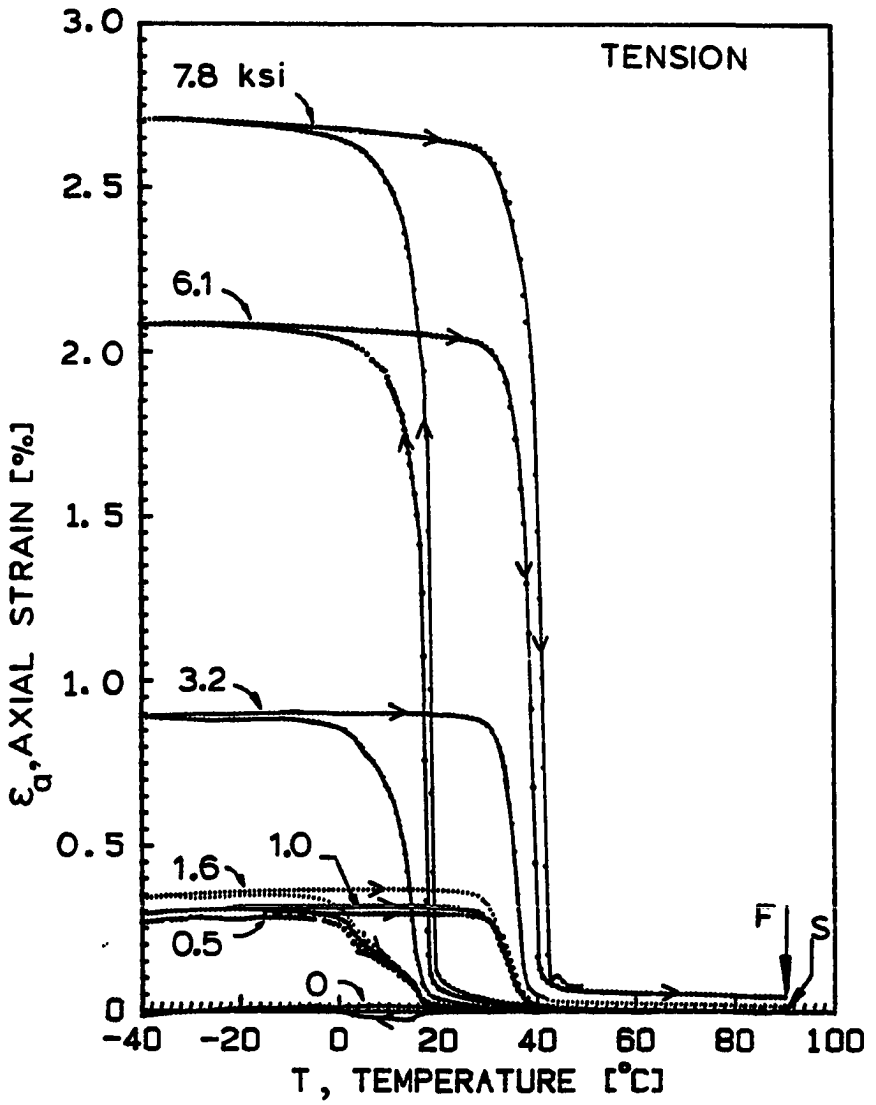


Figure C3. Axial strain (ε_a) versus temperature for various applied uniaxial tensile stresses. Increasing stresses from run to run

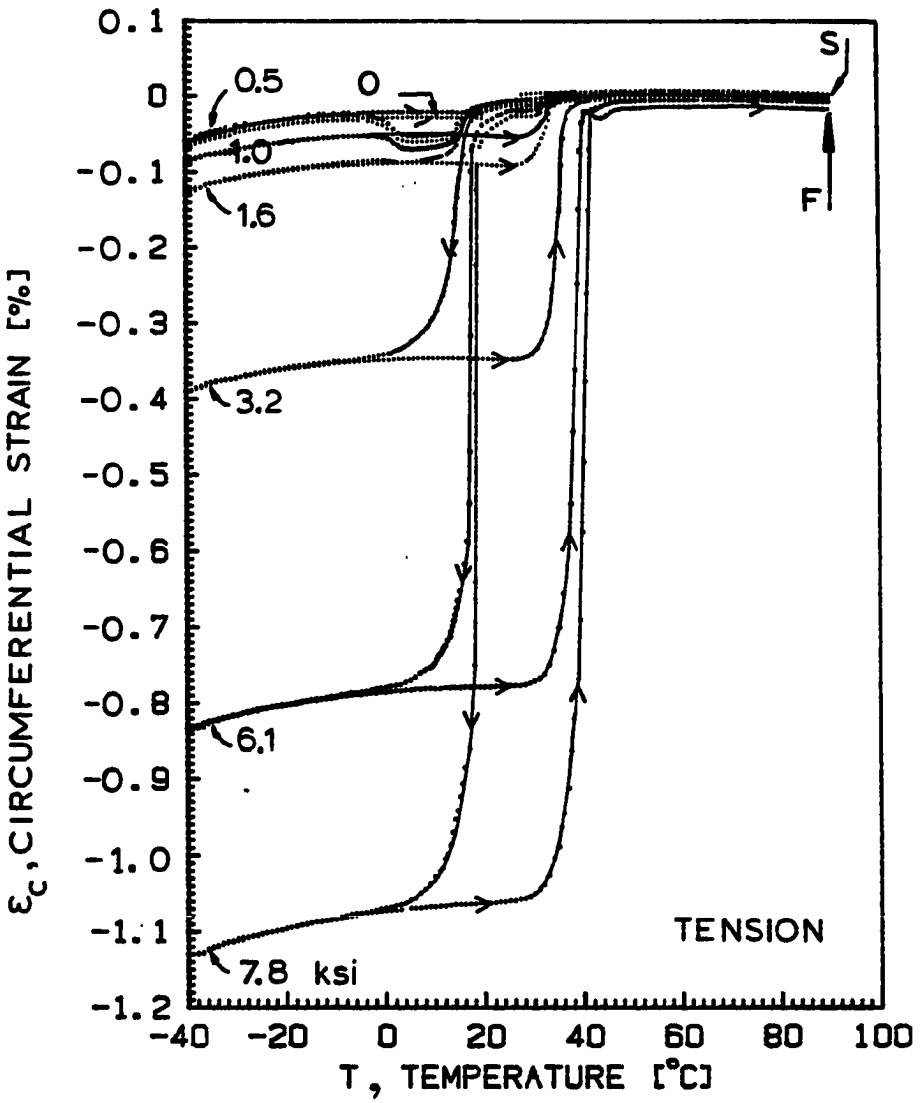


Figure C4. Circumferential strain (ε_c) versus temperature for various applied uniaxial tensile stresses. Increasing stresses from run to run

first runs of compressive stresses in a mixed behavior, as shown in Figs. C5 and C6 for the axial and circumferential strains, respectively. The training can be destroyed also by alternating stresses, even if the direction of them (tensile or compressive) remain the same, as shown in Fig. C7 for the case of alternating zero and increasing tensile stresses compared to the ever increasing tensile stress loading pattern. The effect of preferred activation of variants for the whole regime of tensile and compressive stresses is summarized in Fig. C8.

As a first step in the direction of quantization of the model, the semiaxes of the transformation deformation ellipsoid were determined (Appendix B). In future work, strain field calculations for the martensite-austenite mixture will be carried out. Also, isostrain surfaces will be determined relative to the crystal axes for the different variants. Once this is done, the relative percentage of each kind of variant in each stage of the MPT can be determined for a given applied stress, and hence the anticipated total strain calculated, which can then be compared with the experimental, macroscopic values.

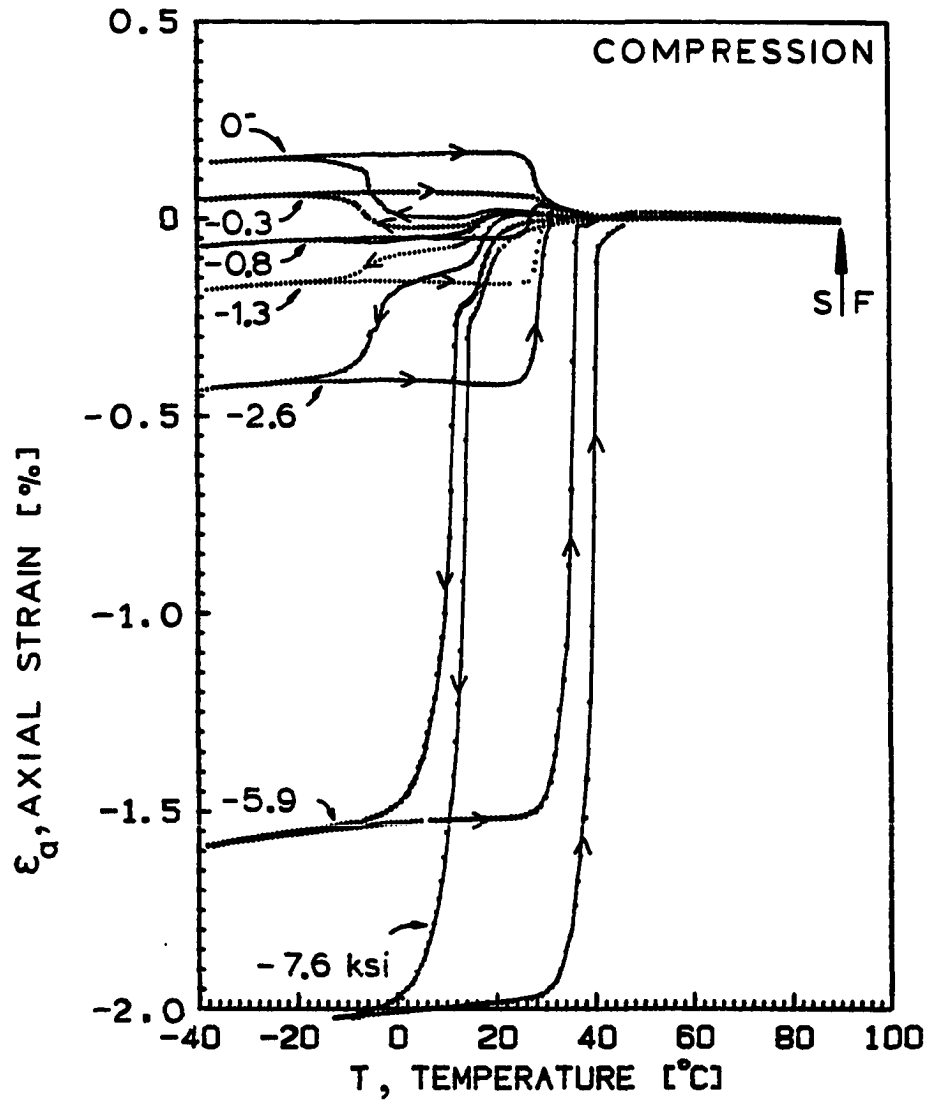


Figure C5. ε_a versus T for various applied uniaxial compressive stresses. Increasing compressive stresses from run to run

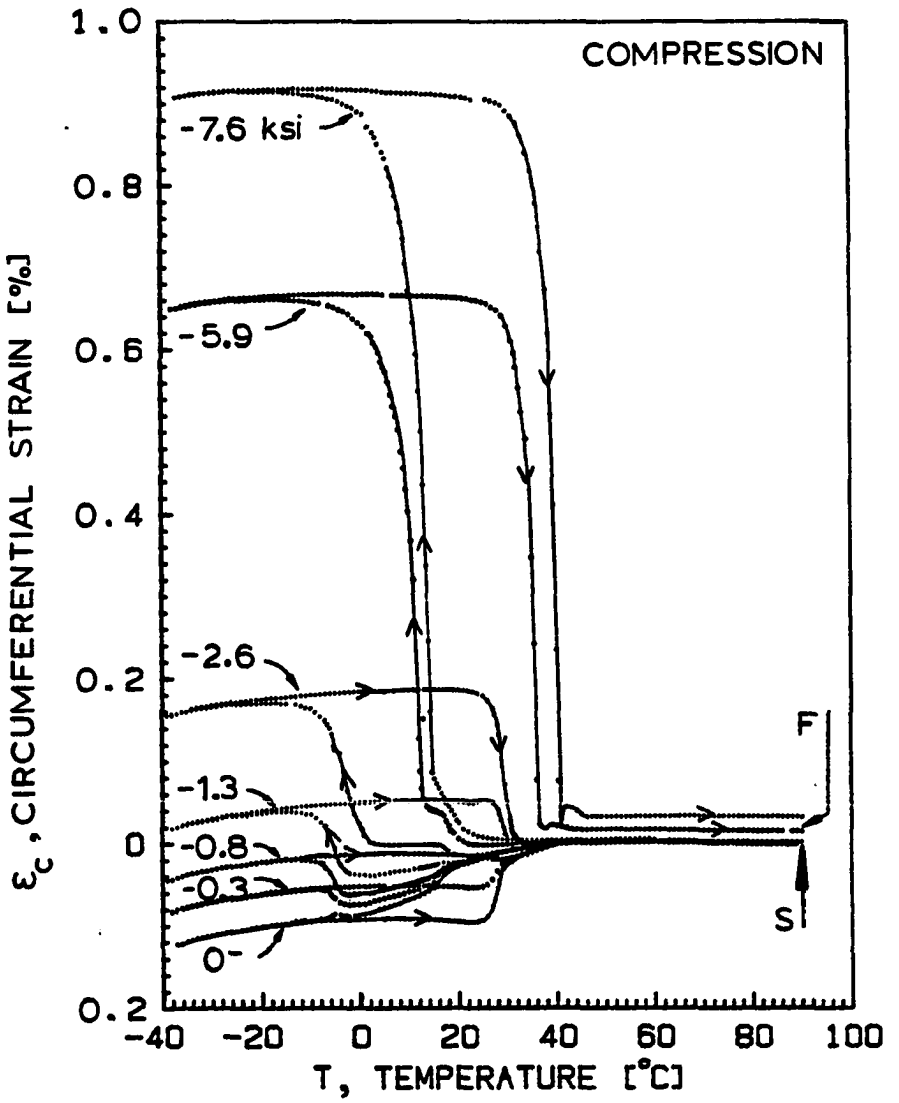


Figure C6. ε_c versus T for various applied uniaxial compressive stresses. Increasing compressive stresses from run to run

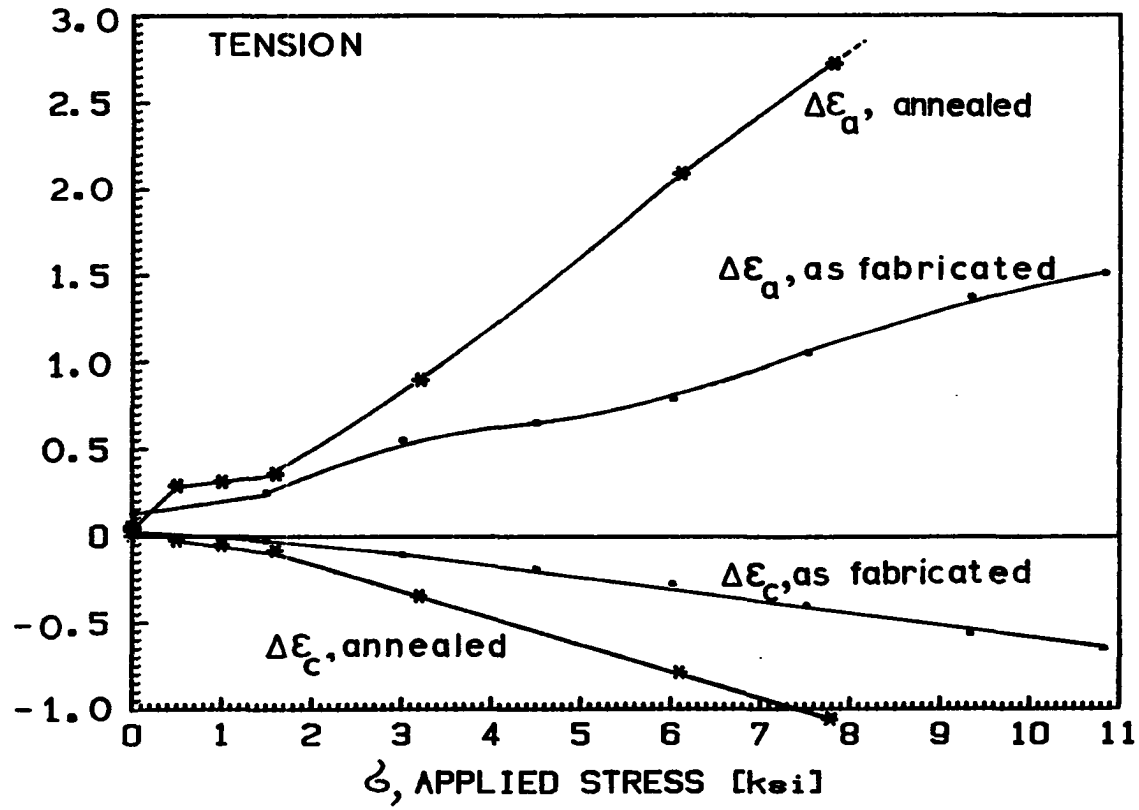


Figure C7. Transformation strains versus applied tensile stress for as-fabricated and annealed sample. For the as-fabricated sample every other run was done at zero stress, whereas for the annealed sample the stress was increased steadily from run to run

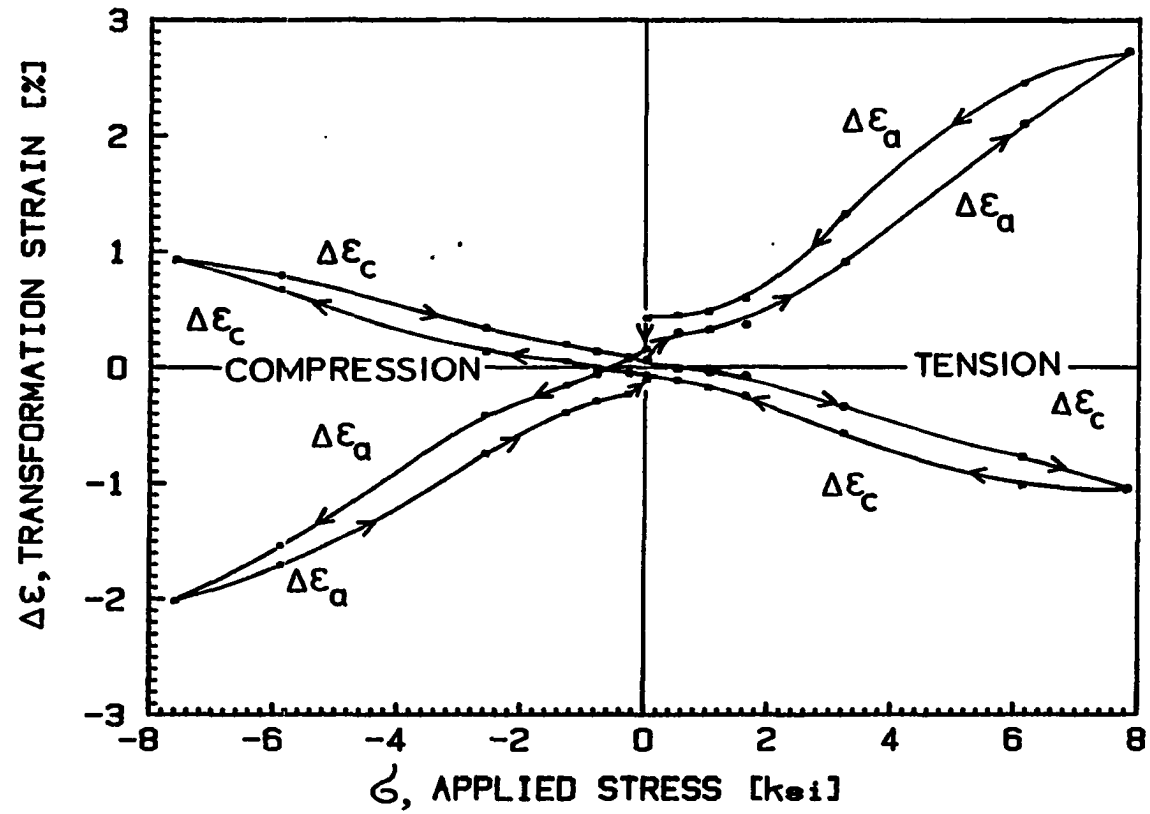


Figure C8. Transformation strains on cooling versus applied stress for axial and circumferential strains and tensile and compressive stresses